A branched-chain 2-oxoacid dehydrogenase multienzyme complex from Thermoplasma acidophilum

Heath, Caroline

Award date:
2006

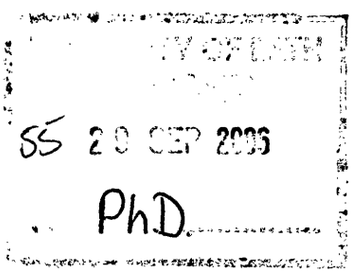
Awarding institution:
University of Bath

A Branched-Chain 2-Oxoacid Dehydrogenase Multienzyme Complex from *Thermoplasma acidophilum*

submitted by
Caroline Heath

for the degree of PhD
of the University of Bath
2006

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346


SS 20 SEP 2006
PhD

Dedicated to my family, and to myself

Acknowledgments

I cannot thank Professor Michael Danson and Dr. David Hough enough, for allowing me the opportunity to carry out research at the University of Bath and for their continuous support throughout. I am honoured to have been chosen by you and to have worked for you.

I should also thank Paul Orsmond, Jim Allpress and Kevin Reiling at Staffordshire University, for help during my bachelor studies and Pauline Gowland, for giving me the opportunity to return to education. Without their support and encouragement in the early stages of my studies, I wouldn't have made it this far. I'd also like to thank Rob Ball for all he did for me at that time.

Thanks to members of labs 1.33 and 1.28, past and present, for technical assistance and many laughs. Particular thanks, for help in the early years of my Ph.D. training, to Drs. Alex Jeffries, Andrew Thompsett, Tina St.John, Mikki Koo, Harry Lamble, Dina Al-Mailam, Ursula Gerike and Nathalie Wery. I am also very grateful to the BBSRC for funding the project.

An eternal thank you to Dina, Tamuna, Shirley, Silvia, Polly, Kim, Amy and Eddie, for tireless moral support throughout the seven years that this career change has taken!

Finally, thanks to my family, for all that they have done for me, and to Lance, my hero, biggest thanks of all.

Abstract

In aerobic bacteria and eukaryotes, the conversion of 2-oxoacids to the corresponding acyl-CoAs in central metabolism is carried out by a family of multienzyme complexes, the 2-oxoacid dehydrogenases (OADHC). The family consists of pyruvate dehydrogenase, 2-oxoglutarate dehydrogenase and branched-chain 2-oxoacid dehydrogenase, which catabolise the conversion of pyruvate, 2-oxoglutarate, and the three branched-chain 2-oxoacids (derived from the transamination of the branched-chain amino acids valine, leucine and isoleucine), respectively.

By contrast, in all the archaea and in anaerobic bacteria, these same metabolic steps are catabolised by a family of 2-oxoacid ferredoxin oxidoreductases (FOR). FORs are much smaller and structurally simpler enzymes than the OADHCs, and the reactions do not involve the same cofactors. It has always been assumed that archaea do not possess OADHCs.

However, in recent years, substantial convincing evidence has emerged to suggest that archaea may in fact have an OADHC. Yet why archaea would utilise the multienzyme complexes, when they employ the simple FOR enzymes, is an intriguing question. Thus, the aim of this project was to establish whether a putative *oadhc* operon found in an archaeal genome does indeed encode a functional OADHC. The archaeon of choice was the euryarchaeote, *Thermoplasma acidophilum*.

As OADHC activity has never been detected in cell extracts of *Tp. acidophilum* (or any other archaeon) with any of the 2-oxoacid substrates, the approach taken was to generate recombinant, OADHC component enzymes. This allowed *in vitro* determination of the substrate specificity and facilitated characterisation of the novel archaeal enzyme.

The research revealed that the putative *oadhc* operon in *Tp. acidophilum* does indeed encode a functional 2-oxoacid dehydrogenase multienzyme complex, and that the complex is a branched-chain 2-oxoacid dehydrogenase (BCOADHC). It has an octahedral 24-mer E2 core and a ~ 165 kDa E1 of $\alpha_2\beta_2$ conformation, like its eukaryotic and bacterial counterparts. Furthermore, kinetic analysis reveals that the branched-chain 2-oxoacids are probably the true substrates for the complex and kinetic values are comparable with eukaryotic and bacterial BCOAHDCs. The *Tp. acidophilum* complex also catalysed the oxidative decarboxylation of pyruvate, but no activity was found with 2-oxoglutarate.

The results suggest that the archaeal BCOADHC should be active *in vivo*, but no activity was detectable in cell extracts of cells grown with either glucose or the branched-chain amino acid precursors to the branched-chain 2-oxoacid substrates. The growth studies revealed that the complex is either not present, or is not active *in vivo*. Thus it cannot yet be said whether archaea actually utilise OADHCs in their central metabolic pathways.

CHAPTER 1: INTRODUCTION

1.1 Archaea and the three Domains of Life

Archaea are one of the three Domains of Life, along with Bacteria and Eukarya. Organisms of Archaea are microscopic prokaryotes and they resemble bacteria in size and shape, which formerly led to their being classed with bacteria (as archaebacteria) in the old five-Kingdom system of classification. However, archaea exhibit vast differences from bacteria at a molecular level, which were revealed through advances in molecular biology (Forterre et al., 2002). The proposal for a new, three-Domain system of classification, to replace the previous system, was made by Carl Woese in 1990, following observations that three, clearly-distinct types of ribosomal RNA existed in different organisms (Woese et al., 1990). This revision led to Archaea being grouped into a Domain of their own. Phylogenetic trees are now routinely compiled based on small subunit ribosomal rRNA sequences (16S or 18S), to classify organisms and show their evolutionary relationships (Fig. 1.1).

Based on 16S rRNA sequences, the Domain Archaea is currently divided into two phyla, Euryarchaeota and Crenarchaeota. Euryarchaeota is the major group and contains almost all the cultivable genera, including the methanogens, the halophiles and the thermophiles. Crenarchaeota contains mainly thermophiles and the hyperthermophiles. A third phylum, Korarchaeota, is also tentatively proposed, based on 16S rRNA sequences obtained from environmental samples (Barns et al., 1996), although these organisms have thus far proved uncultivable. There is disagreement over the authenticity of a fourth phylum, Nanoarchaeota. These organisms have the smallest known genomes (480 kb) and are obligate symbionts of archaea of the genus *Ignicoccus* (Huber et al., 2002). Some researchers propose that the

nanoarchaea in fact belong to Euryarchaeota (Brochier et al., 2005). The majority of characterised archaea are extremophiles, a ubiquitous term used to refer to organisms of any Domain whose optimum growth requirements lie outside those to which man is adapted.

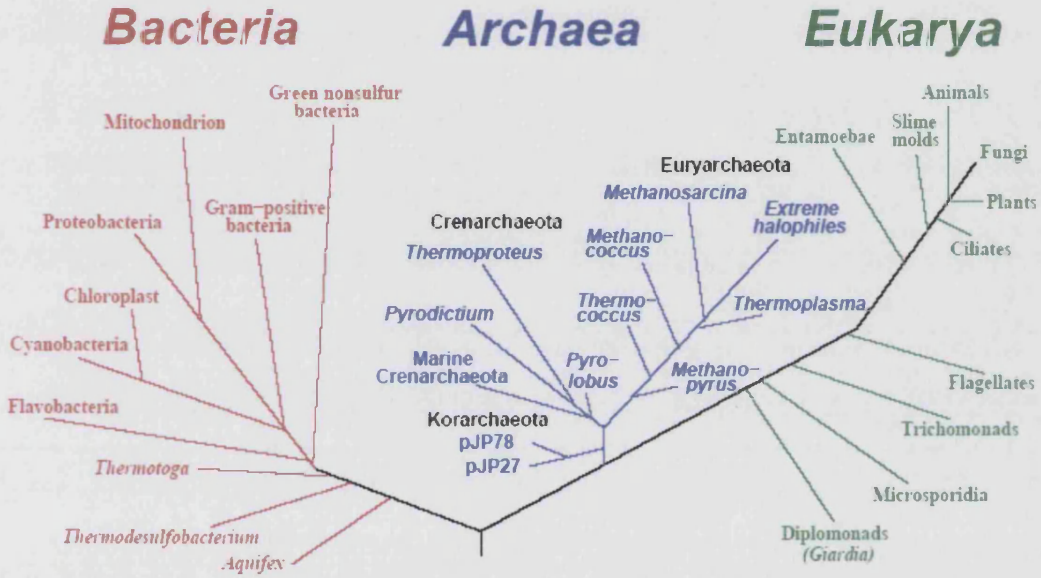


Fig 1.1. Rooted universal phylogenetic tree showing the three domains based upon 16S (or 18S) rRNA sequences. The position of the root was determined by comparing paralogous gene sequences that diverged from each other before the three primary lineages emerged from their common ancestral condition (taken from Jurgens, 2002).

1.2 Central metabolism in Archaea

Many features of central metabolism are common across Archaea, Bacteria and Eukarya. In heterotrophic organisms, central metabolism links the catabolic pathways that degrade organic nutrients to produce energy, and the anabolic pathways of biosynthesis. The central metabolic pathways incorporate those of

glycolysis, the citric acid cycle, and the conversion of pyruvate to acetyl-CoA, which connects the two (Berg et al., 2001) (see Fig. 1.2).

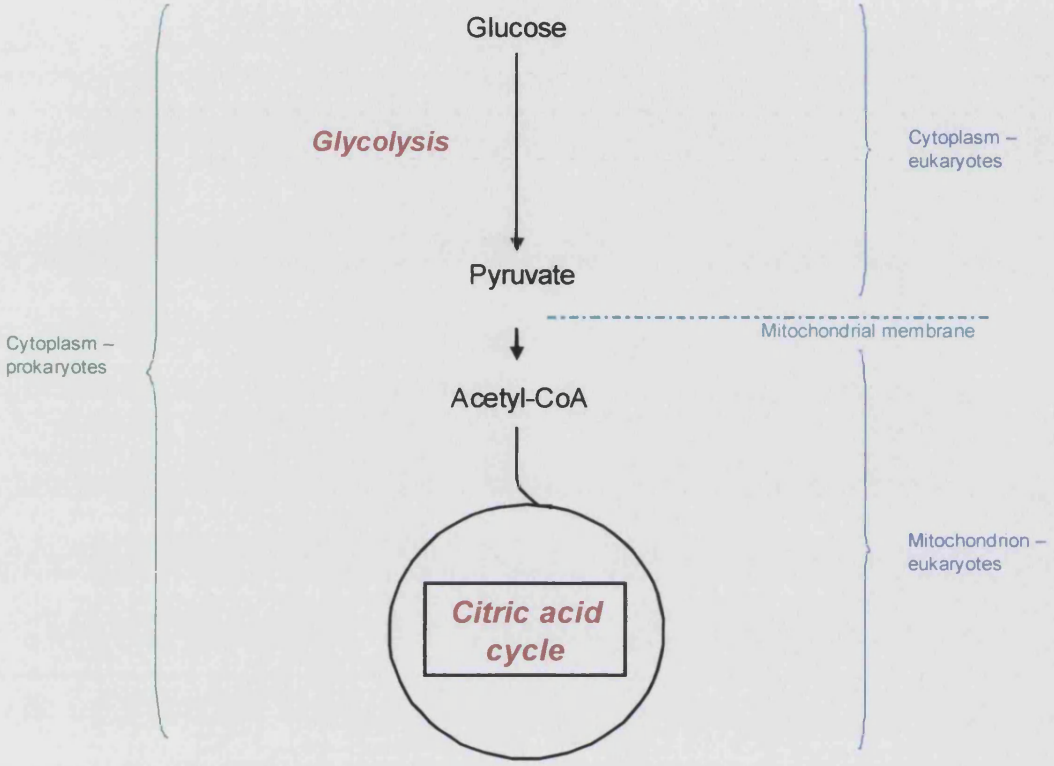


Fig. 1.2. An outline of the pathways of central metabolism, with the cellular location of the processes indicated.

The centrepiece of central metabolism is glycolysis, the conversion of hexose sugars to pyruvate, a process which takes place in the cytoplasm. All organisms, whether aerobic or anaerobic, have the pathway for such glycolytic flux. Bacteria and eukaryotes utilise the Embden Meyerhof (EM) pathway for glycolysis. Archaea utilise either the EM and Entner-Doudoroff pathways, or a variant of them (Danson et al., 2006).

Pyruvate is then converted to acetyl-CoA, which, in aerobes of all Domains, then either feeds into the citric acid cycle and becomes completely oxidised to CO_2,

or is utilised as a precursor in pathways of biosynthesis. In anaerobic organisms, pyruvate serves as an electron acceptor for steps of glycolysis, so that those steps can continue. The citric acid cycle operates at a much reduced rate in anaerobes, although some organisms may possess enzymes of the pathway and it can operate in the reduced mode (Danson et al., 2006).

In eukaryotes, the enzymes involved in the conversion of pyruvate to acetyl-CoA, and in the various steps of the citric acid cycle, are located in the mitochondrion, and pyruvate is transported across the membrane of this organelle. In bacteria and archaea, the degradative processes continue in the cytoplasm.

Some organisms are also able to uptake and catabolise peptides and amino acids, in order to obtain carbon and energy. For example, the first step in the degradation of the branched-chain amino acids valine, leucine and isoleucine is their transamination to branched-chain 2-oxoacids, catalysed by transaminase (aminotransferase) enzymes. The branched-chain 2-oxoacids are then converted to corresponding acyl-CoAs, which similarly feed into the citric acid cycle (again, this process in eukaryotes takes place in the mitochondrion), or serve as precursors in biosythesis (Danson et al., 2006).

1.2.1 Central metabolism in *Thermoplasma acidophilum*

At this point, it is appropriate to mention the pathways of central metabolism in the archaeon *Thermoplasma acidophilum*, as this organism was the subject of this thesis study. *Tp. acidophilum* itself will be introduced later on in the chapter.

Glucose metabolism in *Tp. acidophilum* occurs via a non-phosphorylative Entner-Doudoroff pathway, with no net yield of ATP. Pyruvate is then converted to acetyl-CoA, which feeds into the citric acid cycle. Glucose metabolism of Archaea and the citric acid cycle of *Tp. acidophilum* are shown in Figs. 1.3 and 1.4, respectively.

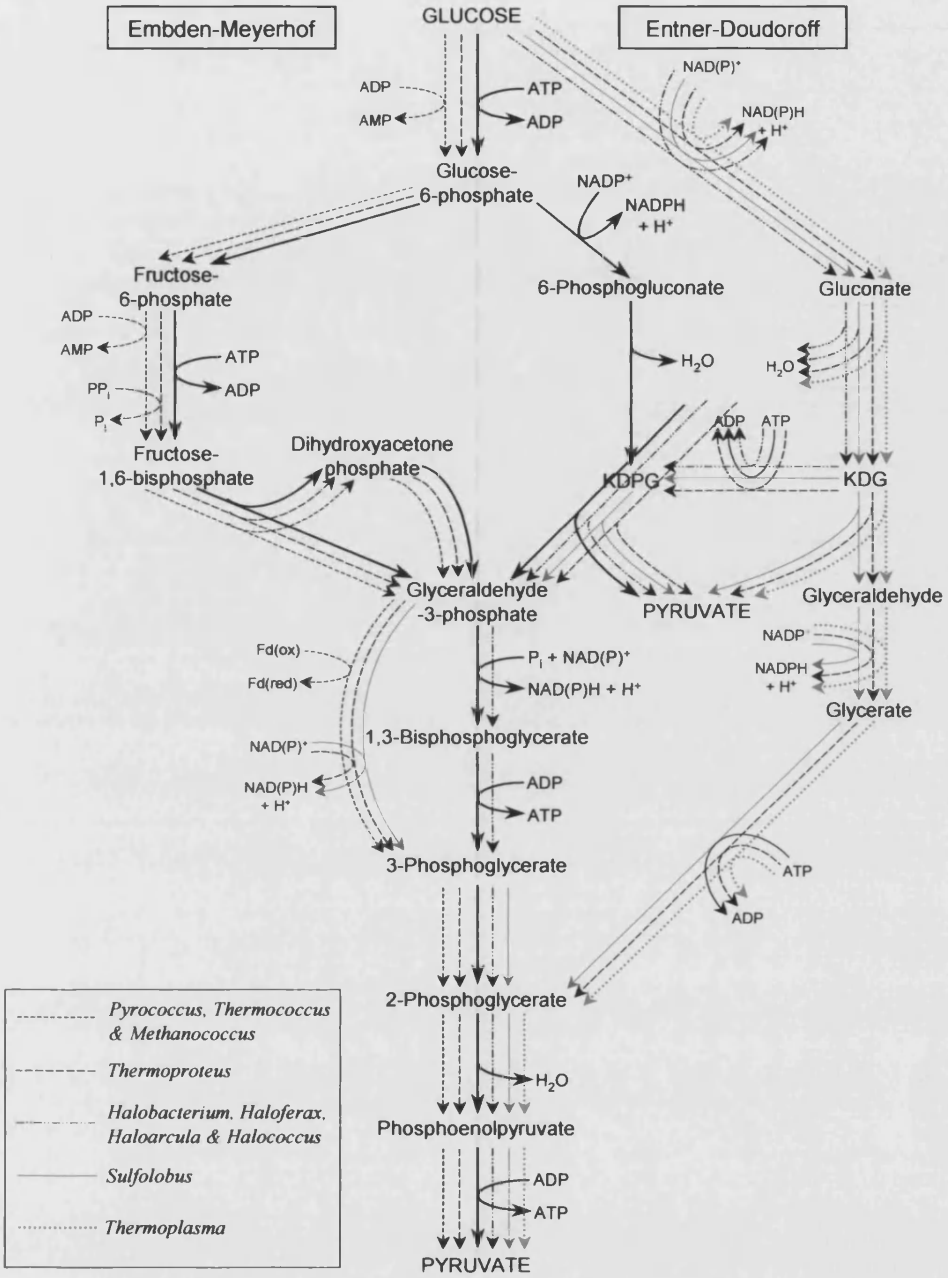


Fig. 1.3. The pathways of glucose metabolism in Archaea. In *Tp. acidophilum*, glucose is converted to pyruvate via the non-phosphorylative Enter-Doudoroff pathway (shown arrowed in red). KDG: 2-keto-3-deoxygluconate; KDPG: 2-keto-3-deoxy-6-phosphate-gluconate.

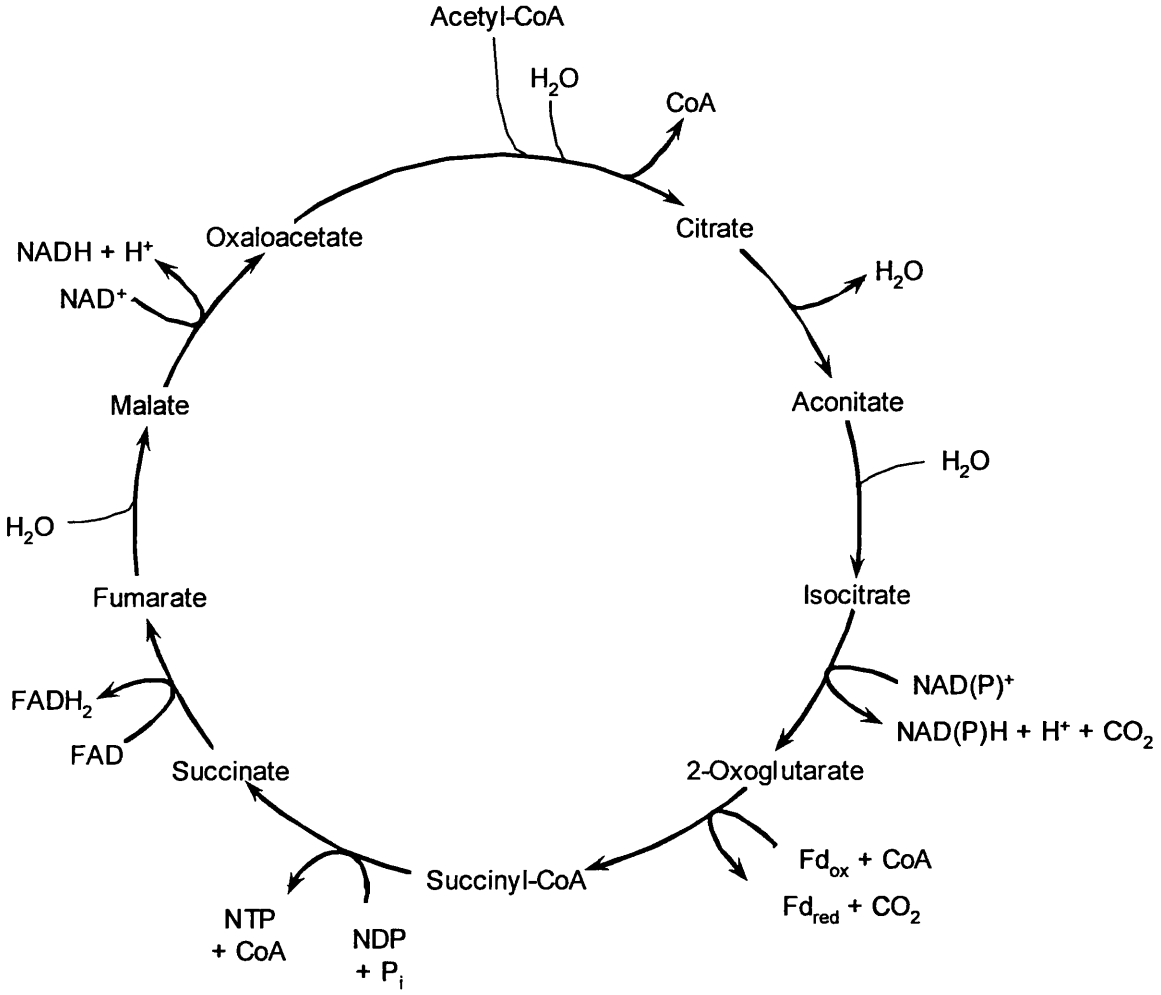


Fig. 1.4. The citric acid cycle of *Tp. acidophilum.*

1.3 The conversion of 2-oxoacids to acyl-CoAs

In all the Archaea, whether aerobic or anaerobic, the conversion of the 2-oxoacids to acyl-CoAs in central metabolism is carried out by a family of enzymes called the 2-oxoacid ferredoxin oxidoreductases (FOR) (Kersher and Oesterhelt, 1981; 1982; Kersher et al., 1982; Schut et al., 2001). Anaerobic bacteria also employ FORs for these reactions. In contrast, the identical step in aerobic bacteria and in eukaryotes is carried out by a family of large multienzyme complexes, the 2-oxoacid dehydrogenases (OADHC) (reviewed in Perham, 1991; 2000).

The reactions catalysed by both enzyme families are thiamine pyrophosphate (TPP) -dependent. However, the two families are structurally very different and they are not homologously related. The FOR family are much smaller, structurally more simple enzymes, and they do not employ the same cofactors as the OADHC family.

The enzymes of the aerobic citric acid cycle (where two of the three members of the OADHC family, the pyruvate and 2-oxoglutarate dehydrogenases, function), are widespread in α-proteobacteria and eukaryotes. Furthermore, the genes encoding these enzymes are homologous, and one evolutionary model proposes that eukaryotes acquired the *oadhc* genes from a bacterial ancestor to the mitochondrion through endosymbiosis (Schnarrenberger and Martin, 2002). This theory would explain the absence of OADHCs in archaea, except that evidence has emerged that members of Archaea may also possess OADHCs. The evidence will be discussed later in the chapter, following a description of the FOR enzymes, and a closer examination of the OADHCs.

1.3.1 2-oxoacid ferredoxin oxidoreductases (FOR)

The FOR family comprises the pyruvate ferredoxin oxidoreductase (POR, EC 1.2.7.1.), which catalyses the conversion of pyruvate to acetyl-CoA and so links glycolysis and the citric acid cycle, the 2-oxoglutarate ferredoxin oxidoreductase

(GOR, EC 1.2.7.3.), which catalyses the conversion of 2-oxoglutarate to succinyl-CoA within the citric acid cycle itself and the 2-oxoisovalerate ferredoxin oxidoreductase (VOR, EC 1.2.7.7.), which oxidatively decarboxylates the branched-chain 2-oxoacids produced by the transamination of valine, leucine and isoleucine (Kersher and Oesterhelt, 1981; 1982; Schut et al., 2001). However, there are examples of FORs demonstrating wider substrate specificity; for example, the VOR of *Thermococcus profundus* has activity with the branched-chain 2-oxoacids and pyruvate, but not 2-oxoglutarate (Ozawa et al., 2005), whilst the FORs of *Sulfolobus* species and *Aeropyrum pernix* are able to utilise both pyruvate and 2-oxoglutarate as substrates (Nishizawana et al., 2005).

Iron sulphur clusters serve as the electron acceptors of the FOR reaction. The acyl-moiety formed on decarboxylation of the 2-oxoacid is handed on directly to CoA, and the reducing equivalents to ferredoxin via the iron-sulphur centre (Kersher and Oesterhelt, 1981), in the following reaction:

$$\text{2-oxoacid} + \text{CoA} + \text{Fd}_{ox} \longrightarrow \text{acyl-CoA} + \text{CO}_2 + \text{Fd}_{red}$$

The catalytic cycle of the FOR comprises a 2-step conversion of the 2-oxoacid to acyl-CoA by thiamine pyrophosphate, and two one-electron transfers. The first step is the TPP-dependent decarboxylation of the 2-oxoacid, and the transfer of one electron per molecule of substrate to the enzyme-bound iron-sulphur cluster. The free-radical intermediate reacts with CoA, causing a second electron to transfer to the iron-sulphur cluster. The acyl group is transferred to CoA, and finally the redox centre is reoxidised (Kersher and Oesterhelt, 1981) (see Fig. 1.5).

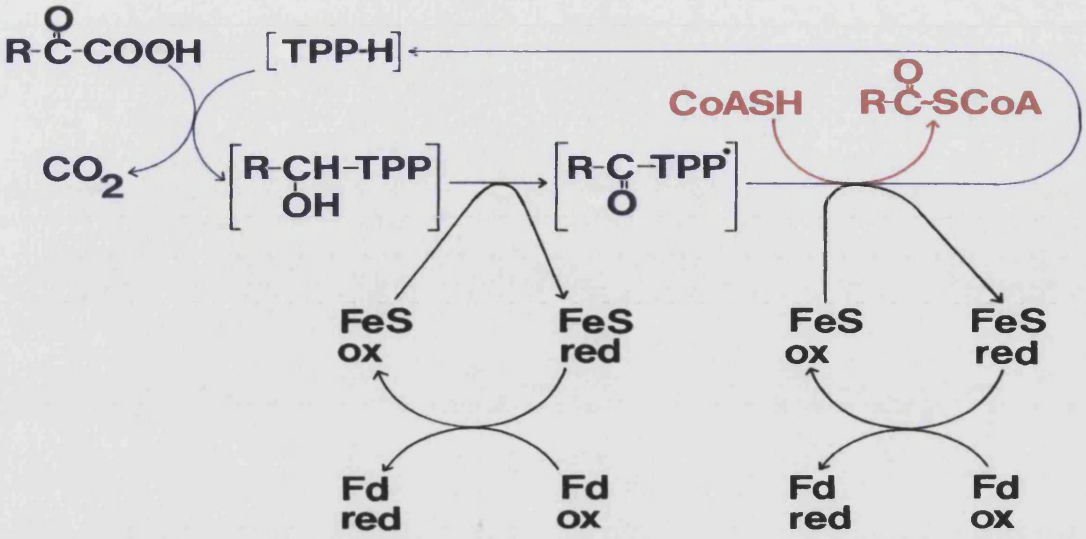


Fig 1.5. An overview of the mechanism of the reactions of 2-oxoacid ferredoxin oxidoreductase of the halophilic archaea (it should be noted that in the methanogenic archaeon, *Methanosarcina barkeri*, the iron-sulphur centres of both the enzyme and the Fd are only reduced in the presence of both substrates, pyruvate and Coenzyme-A and that a similar mechanism may occur in the halophilic enzyme). Fd: ferredoxin; FeS: an enzyme-bound iron-sulphur cluster; TPP: thiamine pyrophosphate (Jolley et al., 2000).

FORs comprise between two and eight polypeptides, and involve combinations of α, β, γ and δ chains, depending on the family type. For example, the POR of *Halobacterium halobium* is an $\alpha_2\beta_2$ structure with a molecular weight of around 256 kDa (Plaga et al., 1992), whilst in the thermophilic archaea FORs generally contain four different subunits and occur as $\alpha_2\beta_2\gamma_2\delta_2$ octamers (Schut et al., 2001). Mesophilic PORs are usually α_2 homodimers (Witzmann and Bisswanger, 1998).

1.3.2 2-oxoacid dehydrogenase multienzyme complexes

In contrast to the relative structural simplicity of the FOR enzymes, the OADHC family comprises enormous enzyme complexes, around 5-10 MDa in size, with diameters of up to 50 nm (Perham, 2000; Perham et al., 2002).

The TPP-dependent reaction catabolised by the OADHC is as follows:

$$\text{2-oxoacid} + \text{CoA} + \text{NAD}^+ \longrightarrow \text{acyl-CoA} + \text{CO}_2 + \text{NADH} + \text{H}^+$$

The OADHC family contains members whose roles are akin to the FORs; members are the pyruvate dehydrogenase complex (PDHC), the 2-oxoglutarate dehydrogenase complex (OGDHC) and the branched-chain 2-oxoacid dehydrogenase complex (BCOADHC), which catabolise the decarboxylation of pyruvate, 2-oxoglutarate and the branched-chain 2-oxoacids, respectively (reviewed by Danson et al., 2004). As with the FOR enzyme family, there are cases where wider substrate specificity is seen in the OADHC; for example, the PDHC of *Bacillus subtilis* utilises both pyruvate and the branched-chain 2-oxoacids, performing a proposed 'dual role' in the processing of these two metabolites (Lowe et al., 1983).

OADHCs comprise multiple copies of three component enzymes: 2-oxoacid decarboxylase (E1) (EC 1.2.4.1, 1.2.4.2 or 1.2.4.4), dihydrolipoyl acyl-transferase (E2) (EC 2.3.1.12 or 2.3.1.61) and dihydrolipoamide dehydrogenase (E3) (EC 1.8.1.4) (Perham, 1991, 2000; Perham et al., 2002, see Fig. 1.6). E1 and E2 are substrate-specific for the particular type of complex to which they belong, but E3 is common to all complex types and catalyses an identical reaction in each (Mattevi et al., 1992). E2 forms the structural core of the complex, with E1 and E3 binding peripherally and non-covalently to the core.

The E2 core molecule comprises multiple polypeptide chains associating into octahedral (24-mer) or icosahedral (60-mer) configurations, depending on the particular complex and the source organism (Reed, 1974; Izard et al., 1999; Perham, 2000). For example, most OGDHCs and BCOADHCs have 24-mer E2 cores, whilst the PDHC of eukaryotes has a 60-mer E2 core (Perham, 2000). E1 may occur as a homodimer or as an $\alpha_2\beta_2$ hetero-tetramer, again depending upon the source and the type of complex (Mattevi et al., 1992). However,

although there are indeed 'trends', there is no fixed pattern of one particular E1 configuration always being associated with a particular E2 core configuration. In all types of complex, though, E3 is a dimer of identical subunits.

1.3.2.1 2-oxoacid decarboxylase

The 2-oxoacid decarboxylase (E1) carries out the first catabolic step in the overall reaction of the OADHC. The two-step reaction involves i) the TPP-dependent oxidative decarboxylation of a 2-oxoacid with formation of a TPP-acyl intermediate and ii) the subsequent acylation of the lipoyl-lysine on the lipoyl domain of E2. The reductive acylation of the lipoic acid moiety of E2, by E1, has been shown to be the rate-limiting step of the overall complex reaction (Bates et al., 1977; Danson et al., 1978, Cate et al., 1980).

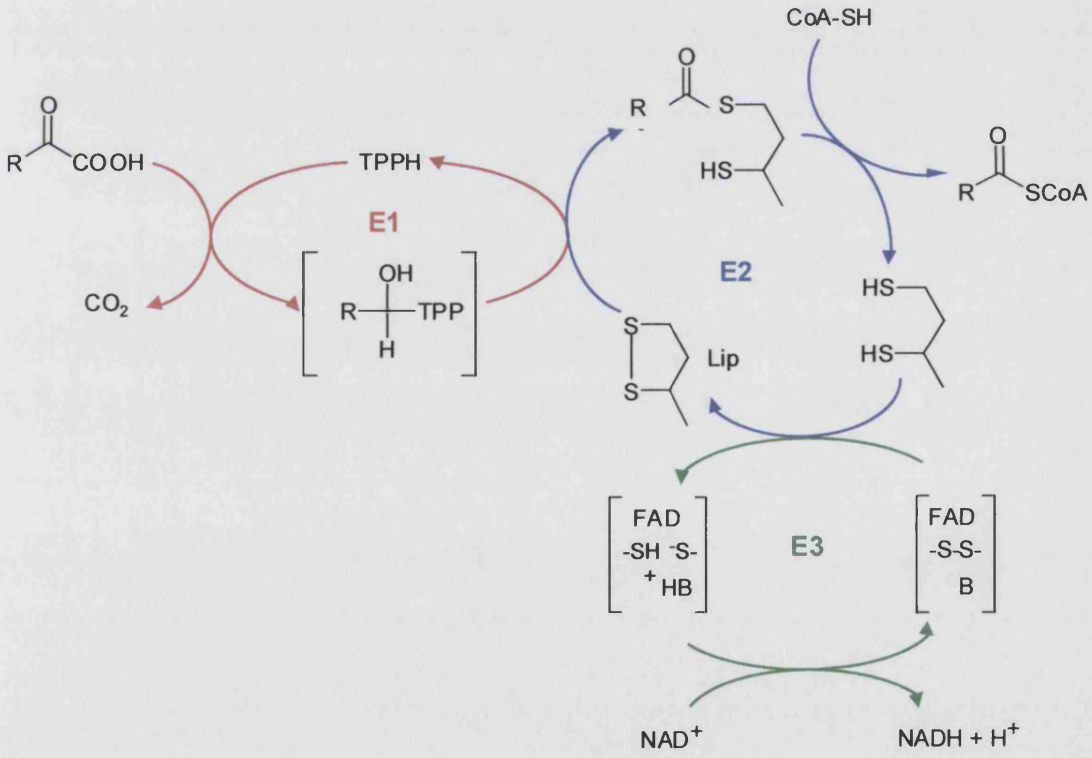


Fig. 1.6. The general reaction catabolised by OADHCs. E1 (2-oxoacid decarboxylase), E2 (dihydrolipoyl acyl-transferase) and E3 (dihydrolipoamide dehydrogenase). Symbols: B (a histidine base on E3); TPP-H (thiamine pyrophosphate); Lip (enzyme-bound lipoic acid).

E1 exists in an α_2 or $\alpha_2\beta_2$ conformation, depending upon the source organism and the type of complex. For example, the E1 of the *E. coli* PDHC is of the α_2 form, whilst the E1 of the *B. stearothermophilus* PDHC is an $\alpha_2\beta_2$ enzyme (Reed et al., 1975; Lessard and Perham, 1994; Arjunan et al., 2002). The molecular weight of a single subunit of the α_2 form is ~ 100 kDa, whilst the α and β subunits of the $\alpha_2\beta_2$ form are ~ 41 and 36 kDa, respectively (Mattevi et al., 1992).

The crystal structure of both forms of E1 has been solved (Aevarsson et al., 1999, Arjunan et al., 2002). For E1 enzymes of $\alpha_2\beta_2$ conformation, for example the 164 kDa E1 component of the BCOADHC of *Pseudomonas putida*, the α_2 subunits act like a 'vice', holding the β subunits in a snug fit, with the C-terminal domains acting as the jaws of the vice, holding on to opposite sides of the β_2 dimer (Aevarsson et al., 1999). E1 $\alpha_2\beta_2$ has two physically remote active sites, located at the interface between the α and β subunits, within a deep funnel-shaped hydrophobic tunnel of around 20 Å in length (Aevarsson et al., 1999). Amino acid residues from both α and β subunits contribute to the active site. The TPP co-factor and a Mg^{2+} ion reside in the active site; TPP is bound to E1α by its phosphate end, and to E1β by its amino-pyrimidine end. Frank et al. (2004) have shown that the two active sites demonstrate conformational asymmetry, and have proposed that the sites communicate via the reversible shuttling of a proton through an acidic tunnel in the protein (termed a 'proton wire'), that allows the sites to switch their conformation alternately and to synchronise events of the reaction that E1 catabolises. The proposed amino acid residues that form the proton wire of the *B. stearothermophilus* PDHC E1 reside on both α and β subunits, and are αAsp 180, αGlu 183, βGlu 28, βGlu 59, βGlu 88 and βAsp 91 (Frank et al., 2004). In E1 $\alpha_2\beta_2$ enzymes, it is the E1β subunit, rather than the E1α subunit, that is responsible for binding of E1 to the E2 core (Wynn et al., 1992).

For E1 enzymes of the α_2 configuration, the two subunits form a dimer with a two-fold axis of symmetry. Each subunit of the dimer has an N-terminal, a middle and a C-terminal domain. The *E. coli* PDHC E1 is a dimer of 200 kDa (Arjunan et al., 2002). Like E1 $\alpha_2\beta_2$, the enzyme has two active sites, located deep within a 21-Å funnel-shaped hole at the dimer interface. The TPP cofactor and Mg^{2+} ion are situated in the active site; the pyrophosphate end of TPP binds to N-terminal residues of one subunit, whilst the amino-pyrimidine end binds to both the same N-terminal domain and to the middle domain of the second subunit. The N-terminal domain of E1 is responsible for binding to the E2 core. The interaction is thought to be electrostatic in nature, involving negatively-charged residues on E1 and positively-charged ones on E2 (Hengeveld et al., 1999; Arjunan et al., 2002)

There are structural similarities between the domains of E1α_2 and E1$\alpha_2\beta_2$ enzymes. The N-terminus of E1 α_2 corresponds to the α subunit of $\alpha_2\beta_2$, the middle domain of α_2 corresponds to the N-terminus of $\alpha_2\beta_2$-E1β and the α_2 C-terminal domain to the C-terminus of $\alpha_2\beta_2$-E1β. However, there are also several differences, such as the orientation of some amino acid residues in the different subunits, the absence of certain helices from E1$\alpha_2\beta_2$, and differences in the length of loop regions between particular β-strands (Arjunan et al., 2002).

E1 TTP-binding motif

Hawkins et al. (1989) describe a common sequence motif found on the E1α of OADHCs, and propose that the motif is associated with TPP binding. The region is generally around 30 amino acids in length, with highly conserved GDG and NN residues at the N- and C- termini respectively. Internally, there is an E or D residue that is usually conserved, and A and P residues that are generally conserved. A cluster of hydrophobic residues immediately precedes the NN (Hawkins et al., 1989). This motif has also been identified in the amino acid sequence of the putative E1α of the archaeon *Haloferax volcanii* (Jolley et al., 2000).

Regulation by reversible phosphorylation

The activity of some enzymes may be regulated by reversible phosphorylation. In this method of regulation, target proteins are continuously inactivated (phosphorylated) by intrinsic protein kinases, and reactivated (dephosphorylated) by phosphatases (although for some enzymes, the reverse effect takes place). Some OADHCs undergo reversible phosphorylation, and the covalent modification occurs at specific serine residues on the E1α subunit. These serine residues are conserved in numerous eukaryotic, bacterial and putative archaeal E1α subunits, and are highlighted in Fries et al. (2003). The mammalian BCOADHC and PDHC have both been shown to be regulated by reversible phosphorylation (Paxton and Harris, 1982; Cook et al., 1984; Wynn et al., 2004; Holness and Sugden, 2003), and recently it has been suggested that some metabolic enzymes of the archaeon *Sulfolobus solfataricus* may undergo reversible phosphorylation as a method of regulation (Ray et al., 2005; Kim and Lee, 2005).

1.3.2.2 Dihydrolipoyl acyl-transferase

Dihydrolipoyl acyl-transferase (E2) catalyses the acyltransferase reaction, and also forms the highly symmetrical structural core of the OADHC. The enzyme is able to couple the reactions of the individual subunits of the complex by means of an enzyme-bound lipoyl group that visits the active sites of the component enzymes (Mattevi et al., 1992).

E2 comprises between one and three N-terminal lipoyl domains, a peripheral subunit binding domain (PSBD) and a C-terminal catalytic domain (Mattevi et al., 1992). The three domains are interconnected by proline- and alanine-rich semi-flexible linker regions, some 20-30 amino acid residues in length.

The lipoyl domain

The lipoyl domain is around 80 amino acid residues in length. The domain undergoes lipoylation, a post-translational modification involving covalent

attachment of a lipoyl moiety to a specific lysine residue. The domain serves as a swinging arm that connects the active sites of E1, E2 and E3, thus facilitating substrate channelling through the complex (Perham et al., 2002). It is the lipoyl domain, rather than the lipoic acid moiety alone, that is the true substrate recognised by the E1, E2 and E3 active sites (Perham et al., 2002). Rather than being free-swinging, the lipoyl-lysine residue appears to have a preferred orientation. This feature is thought to facilitate insertion of the lipoyl-lysine arm into the E1 active site, as the arm needs to be fully extended and in a particular orientation, in order to reach into the deep funnel in which the active site is located (Jones et al., 2000; Perham et al., 2002). Free lipoic acid is, in fact, a relatively poor substrate, and so recognition of the lipoyl domain by the active sites of the complex provides an elegant mechanism for substrate channelling in this system.

Organisms may synthesise lipoic acid endogenously, or scavenge it from the environment (Miller et al., 2000 and references therein). *E. coli* employs two discreet lipoylation pathways in order to achieve lipoylation of E2. The endogenous pathway utilises the enzymes lipoyl synthase (LipA) and lipoyltransferase (LipB). Acyl-carrier protein-bound octanoic acid, an intermediate of fatty acid biosynthesis, is covalently attached to the target lysine residue of E2 by LipB. LipA then adds two sulphur atoms, resulting in lipoylated E2 protein. The exogenous pathway employs a lipoate protein ligase (lplA). The pathway involves uptake of lipoic acid from the external environment and transport across the cell membrane into the cell. It is then covalently attached to the lysine residue in the lipoyl domain of E2 by lplA, in an ATP-dependent reaction (Morris et al., 1994, 1995; Miller et al., 2000) (see figs. 1.7 and 1.8).

The three dimensional structure of the lipoyl domain from a wide range of sources has been determined by NMR spectroscopy (Perham, 2000, and references therein); all the domains show sequence similarity and adopt a similar three-dimensional structure. The domain comprises a β-barrel, made up

of two, four-stranded β-sheets (β-strands 1,3,6,8 and 2,4,5,7). The lysine residue that is the target for lipoylation occupies an exposed position in a surface loop region connecting β-strands 1 and 2 (Reche and Perham, 1999) and in many cases is flanked by moderately conserved D and A/V residues. The lipoylation process can take place, with varying efficiencies, across the species barrier, as it is the exact positioning of the lysine in the β-turn of the domain, rather than the exact identity of the neighbouring residues, that is fundamental to target lysine recognition by the ligase enzyme. If situated just one residue towards the N or C termini of the polypeptide, lipoylation does not occur (Wallis and Perham, 1994).

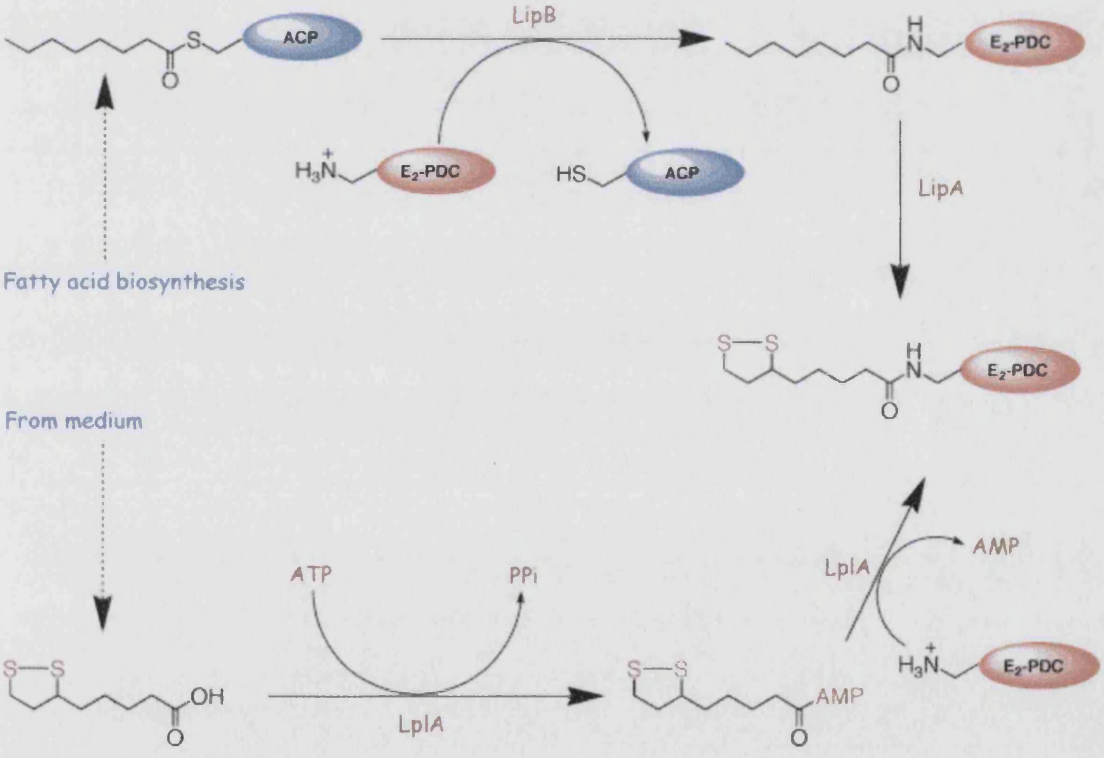


Fig. 1.7. The two pathways of lipoylation in *E. coli*: the endogenous method, which employs LipA and LipB, and the exogenous method, which employs lplA.

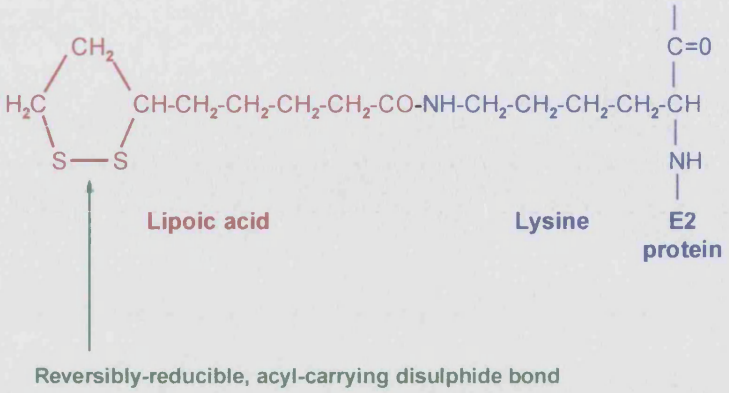


Fig. 1.8. Lipoic acid, a sulfur-containing cofactor, covalently attached to the side chain of the target lysine residue in E2.

Peripheral subunit binding domain

Binding of E1 and E3 to the E2 core occurs in the peripheral subunit binding domain (PSBD). Given its function, this is a remarkably small domain of around 35 amino acid residues. It lies between the N-terminal lipoyl and C-terminal catalytic domains. In the octahedral BCOADHC E2 and the icosahedral PDHC E2, the PSBD binds E1 or E3 mutually exclusively. Thus, spatial competition between E1 and E3, for binding to the PSBD, is witnessed (Lessard et al., 1998; Perham, 2000). The situation is slightly different in octahedral OGDHC and PDHC E2s, as E1 also binds in part to the catalytic domain (Lessard et al., 1998, and references therein).

The catalytic domain

The active site of E2 resides in the catalytic domain, and lies within a long channel at the interface of the E2 subunits. E2 trimers form 'building blocks' that then assemble into the symmetrical core comprising 24 or 60 polypeptides with cubic or dodecahedric shape, respectively. The core molecule has 'windows' in each face, the purpose of which appears to be to allow the passage of substrate and product to and from the E2 active site from within the centre of the core. E1

and E3, bound around the periphery, form a shell that encases the core. This outer shell is separated from the core by an annular space, within which the lipoyl domain can freely access the active sites of all three component enzymes (Mattevi et al., 1992; Milne et al., 2002; Izard et al., 1999).

Bates et al. (1977) demonstrated that acyl groups can be transferred between lipoyl moieties bound to neighbouring E2 chains of the core (termed 'transacetylation') and the process has the potential for a single E1 enzyme to supply acyl groups to around 12 localised E2 polypeptides (Bates et al., 1977). This reaction occurs faster than the initial reductive acetylation reaction between E1 and E2. Transacetylation is thought to assist in preventing diffusion of substrate in an environment low in 2-oxoacid and CoA concentrations, increasing the likelihood of contact between the CoA substrate and the lipoyl-bound acyl group and thus enhancing the overall complex reaction rate (Danson et al., 1978). The process is also referred to as 'active site coupling'.

1.3.2.3 Dihydrolipoamide dehydrogenase

Dihydrolipoamide dehydrogenase (E3) carries out the FAD-dependent reoxidation of the E2 dihydrolipoyl moiety. Like E1, E3 catalyses two consecutive half-reactions. The first is the transfer of two electrons from E2 lipoamide to enzyme-bound FAD and two sulphur atoms in the active site, and the second is the transfer of the electrons to NAD^+ and the return of the enzyme to its oxidized starting state (Mattevi et al., 1992).

E3 is an α_2 homodimeric enzyme, with an overall subunit molecular weight of around 110 kDa. Its structure has been solved (Mattevi et al., 1991) and the enzyme was shown to adopt a characteristic butterfly-like conformation. Each polypeptide subunit has four domains: an FAD-binding domain, an NAD-binding domain, a central domain and an interface domain. The active site resides at the interface of the two subunits, with one molecule of FAD bound to each. The redox centre of the active site comprises a flavin ring and an adjacent disulphide

bridge, and functions in the transfer of the reducing equivalents from dihydrolipoamide to NAD^+. The binding sites for dihydrolipoamide and NAD^+ are on opposite faces of the flavin cofactor (Mattevi et al., 1992). Conserved residues include a G-X-G-X-X-G motif responsible for cofactor binding, a conserved motif surrounding an active-site histidine, and two active-site cysteine residues separated by four amino acids (Vettakkorumakankav and Stevenson, 1992, and references therein).

1.4 Evidence for OADHCs in Archaea

Archaea are thought to utilise exclusively the FOR to catalyse conversion of 2-oxoacids to their corresponding acyl-CoAs, and it has always been assumed that OADHCs are not present in these organisms (Kersher and Oesterhelt, 1981; 1982; Schut et al., 2001). However, there is now increasing evidence that OADHCs may indeed exist in the archaea (reviewed in Danson et al., 2004).

Danson et al. (1984; 1986) reported the first findings of archaeal DHlipDH activity, in the halophilic archaeon *Halobacterium halobium.* The only known function of this enzyme to date is as the E3 component of an OADHC, or of the glycine cleavage system, neither of which were thought to be present in archaea. E3 activity has subsequently been detected in *Thermoplasma acidophilum* (Smith et al., 1987). Following the discovery of E3, the lipoic acid substrate, which serves as acyl-carrying cofactor in OADHCs, was detected in *Halobacterium halobium*, using a covalent chromatography and GC-mass spectrometric procedure (Pratt et al., 1989).

E3 of *Haloferax volcanii* was cloned and sequenced using information obtained through purification of the native protein (Vettakkorumakankav et al., 1992; Vettakkorumakankav and Stevenson, 1992). The amino acid sequence, derived from the acquired gene sequence, possessed a high degree of homology when aligned with the DHlipDH of the OADHCs from eukaryotic and bacterial sources. Conserved residues involved in the active site and in cofactor binding were also

identified. Jolley et al. (2000) sequenced the region upstream of the DHlipDH gene of *Hfx. volcanii,* and it became apparent that the gene was the 4th open reading frame in a gene cluster. BLAST searches revealed that the protein products of the other 3 putative ORFs of the operon showed homology to the E1α, E1β and E2 enzymes of the PDHC from the Gram-positive bacterium *Bacillus stearothermophilus.* The archaeal ORFs are tightly spaced and lie in the same orientation to each other. There are no obvious promoters or Shine-Dalgarno sequences upstream of ORFs 2, 3 or 4. However, ~40 bp upstream of ORF 1 there are two AT-rich regions, either of which could serve as promoter, and 8 bp upstream of the initiation codon of ORF1 there is a strong Shine-Dalgarno sequence (CAGGAGG). Downstream of ORF 4 is a good transcriptional stop signal (poly-dT tract) (Jolley et al., 2000).

E2 serves as both the catalytic and structural core to the complex, and has distinct domains related to its various functions. Comparisons of the predicted secondary structures of the archaeal E2 polypeptides from *Hfx. volcanii* with that of the E2 of *B. stearothermophilus* revealed almost identical structural domain features and conserved amino acid regions.

Jolley et al. (2000) also demonstrated that the putative OADHC operon of *Hfx. volcanii* may be transcribed as a single message. The total RNA from *Hfx. volcanii* was resolved by agarose gel electrophoresis and subjected to Northern Analysis, using the vector pNAT82, containing the cloned region of E3. A 5.2 kb RNA species hybridized to pNAT82, a length very close to the 5.4 kb predicted length of the putative operon.

As increasing numbers of archaeal genomes have been sequenced, it has been possible to search for putative 2-oxoacid dehydrogenase complex operons. Indeed, in aerobic halophilic and thermophilic species, putative operons have been identified, with 4 open reading frames arranged with a potential Shine-Dalgarno sequence and a TATA promoter sequence upstream of the start codon

of ORF 1, and a poly-dT tract downstream of the stop codon of ORF 4 (Fig. 1.9) (Danson et al., 2006). Conserved residues involved in active-site and in cofactor binding have been identified, by sequence alignment, in the archaeal species *Tp. acidophilum*, *Aeropyrum pernix*, *Haloferax volcanii* and *Sulfolobus solfataricus*. Secondary structural predictions of the E2 of *Tp. acidophilum* have also been performed, and are shown in Fig. 1.10. (Morgan, 2001).

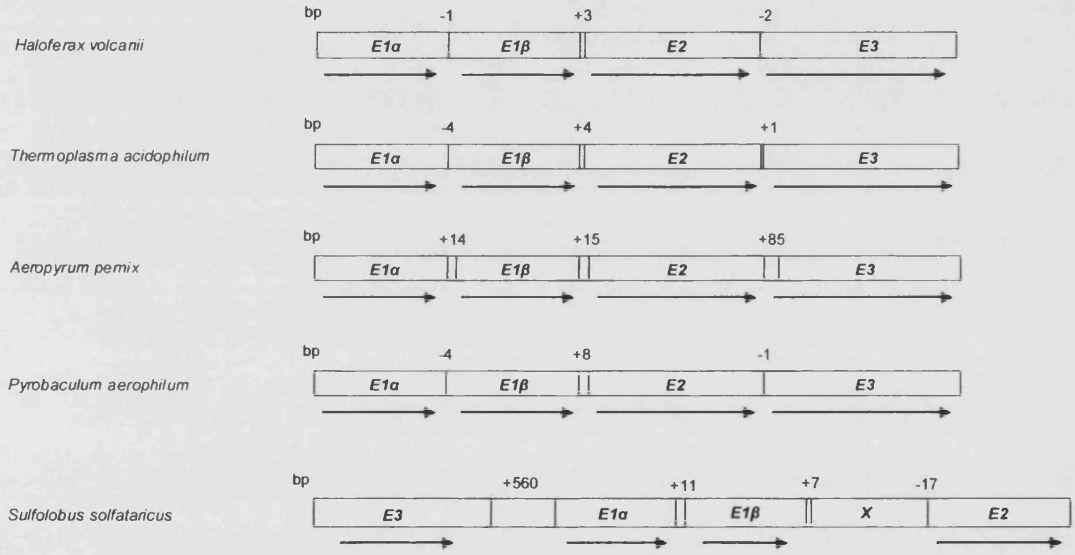


Fig 1.9: Genome analyses. The arrangement, intergene distances (bp) and proposed directions of transcription of the open reading frames (ORFs) constituting the E1α, E1β, E2 and E3 genes of the proposed 2-oxoacid dehydrogenase complex operons of *Hfx. volcanii*, *Tp. acidophilum*, *A. pernix, P. aerophilum* and *S. solfataricus*. The protein X gene from *S. solfataricus* is of unknown function, but closely matches protein AcoX of the acetoin dehydrogenase operon of *Alcaligenes eutrophus* (Danson et al., 2006).

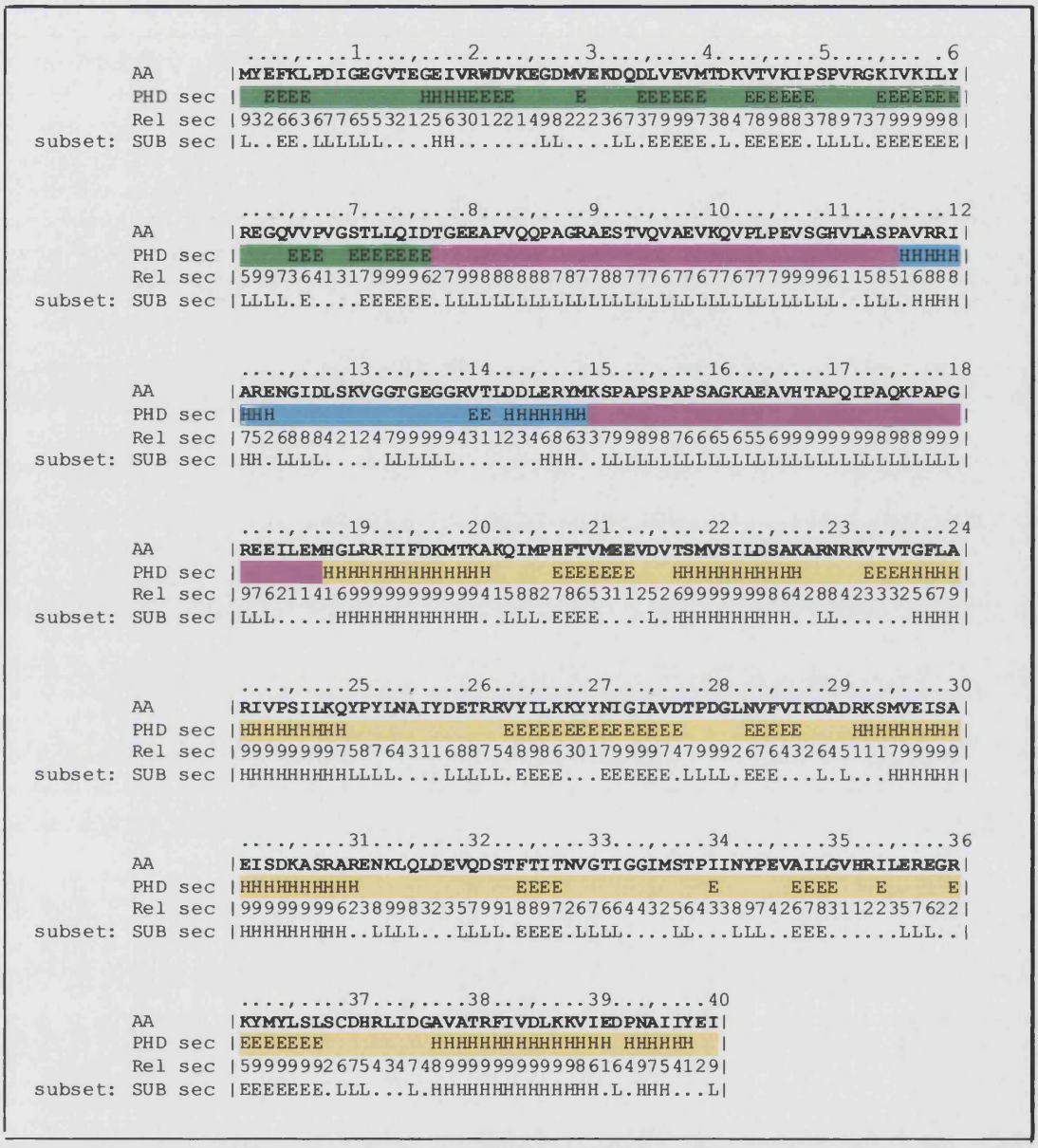


Fig 1.10. Secondary structure predictions of the *Tp. acidophilum* E2. The primary amino acid sequence is highlighted in bold and numbered in tens of residues. 'PHD sec' is the predicted secondary structure at each residue (H is predicted α-helix, E is predicted extended β-sheet and spaces represent predicted loop regions). 'Rel sec' is the reliability out of ten for the structural prediction at each amino acid position. 'SUB sec' is residues where the structural prediction reliability is greater than 80 %. Colouring donates structural domains of the E2 protein; green is the lipoyl domain, blue is the PSBD and yellow is the catalytic and structural core domain. Flexible linker regions are shown in magenta (from Morgan, 2001).

1.5 An OADHC in *Thermoplasma acidophilum*?

Thermoplasma acidophilum is a thermoacidophilic archaeon belonging to the phylum Euryarchaeota. The organism is a facultative anaerobe and has growth requirements of 45-62 °C and pH 0.9-3.5, with optimum growth occurring at 59 °C and pH 1-2. It has been isolated from coal refuse piles and solfatara fields (Darland et al., 1970; Segerer et al., 1988, Yasuda et al., 1995).

The organism has a small (1.56 Mb) genome, which has been fully sequenced (Ruepp et al., 2000). This archaeon is one of those whose genome contains a putative *oadhc* gene cluster (Morgan, 2001). It is also one of the organisms in which activity of the E3 enzyme has been detected (Smith et al., 1987) and the E3 gene has been cloned and expressed as a soluble, active protein in *E. coli* (Aass, 2003). However, OADHC activity has never been detected in cell extracts of *Tp. acidophilum*, or any other archaeon (Danson et al., 2004). *Tp. acidophilum* also possesses at least one *for* operon that would encode an FOR, although activity of this enzyme with any 2-oxoacid substrate has not, to date, been confirmed by enzymological studies.

1.6 Aims

The evidence strongly suggests that aerobic members of Archaea do possess OADHCs, but why would archaea need these multienzyme complexes when they employ more simple FOR enzymes that catabolise the identical metabolic steps? The first step in answering this question would be to establish whether a putative *oadhc* operon found in an archaeal genome does actually encode a functional OADHC.

The aim of the project was to determine whether the putative *oadhc* of *Tp. acidophilum* encoded a functional OADHC. As OADHC activity has never been detected in cell extracts of the organism with any of the 2-oxoacid substrates, the approach taken was to generate recombinant, OADHC component enzymes. This would facilitate *in vitro* determination of the substrate specificity and allow characterisation of this novel archaeal enzyme.

CHAPTER 2: MATERIALS AND METHODS

2.1 General laboratory reagents

In all cases, unless otherwise stated, laboratory reagents and bacteriological media were supplied by Sigma-Aldrich Company Ltd. (Poole, UK), Difco Laboratories (Detroit, USA) and Fisher Scientific Ltd. (Loughborough, UK). DNA polymerases and restriction enzymes were from New England Biolabs Ltd. (Hitchin, UK) or Promega (Southampton, UK). Primers were manufactured by MWG Biotech (Ebersberg, Germany). dNTPs were supplied by Bioline (London, UK). pET expression vectors, *E. coli* expression strains, BugBuster Protein Extraction Reagent and benzonase nuclease were purchased from Novagen-Merck (Nottingham, UK). *E. coli* JM109 cells, pGEM-T vector, restriction endonucleases, T4 DNA ligase and *Taq* DNA polymerase were purchased from Promega. Vent DNA polymerase was from New England Biolabs (Hitchin, UK). Lipoic acid, phenylmethanesulfonyl fluoride (PMSF) and antibiotics were purchased from Sigma. Molecular mass markers were from Bio-Rad (Hemel Hempstead, UK) or Invitrogen (Paisley, UK).

2.2 Microbiology techniques

2.2.1 Growth of *Escherichia coli*

Strains were cultured in liquid Luria Bertani (LB) medium adjusted to pH 7.0 with NaOH, containing sodium chloride (1 % w/v), tryptone (1 % w/v) and yeast extract (0.5 % w/v), or on LB agar (1.5 % w/v agar) plates. Media were supplemented with antibiotic when required (carbenicillin, 50 µg/ml; kanamycin, 30 µg/ml; chloramphenicol, 34 µg/ml) as recommended in the Novagen pET System manual (9th edition) for pET vectors, or the Promega manual (1999) for pGEM-T vectors.

2.2.2 Growth of *Thermoplasma acidophilum*

Growth of *Tp. acidophilum* is described fully in Chapter 6.

2.3 Molecular biology techniques

2.3.1 Preparation of *Escherichia coli* JM109 competent cells

A stock of competent cells was prepared following the Promega Protocols and Applications Guide (3rd edition). A 250-ml culture, in LB media supplemented with 20 mM $MgSO_4$, was grown to mid-exponential phase in conical flasks and then harvested by centrifugation at 4500 × *g* for 5 min at 4 °C. Steps were then carried out on ice, using chilled apparatus. Cells were resuspended in 0.4 × original volume of ice-cold TFB1 (30 mM CH_3COOK, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 100 mM RbCl, 15 % [v/v] glycerol, adjusted to pH 5.8 with 1 M acetic acid), incubated on ice for 5 min, and harvested by centrifugation as before. Cells were then resuspended in 0.25x the original volume of ice-cold TFB2 (10 mM MOPS or PIPES, 75 mM $CaCl_2$, 10 mM RbCl, 15 % (v/v) glycerol, adjusted to pH 6.5 with 1 M KOH) and incubated on ice for 20 min. Competent cells were aliquoted (200 µl), snap-frozen in dry ice / isopropanol and stored at -80 °C.

2.3.2 Preparation of competent cells using $CaCl_2$

LB medium (plus antibiotic when required) was inoculated with ~ 5 colonies and incubated at 37 °C, 200 rpm until A_{600} = 0.6. *E. coli* cells were pelleted by centrifugation at 3000 × *g* for 5 min. The pellet was resuspended in sterile, ice-cold 50 mM $CaCl_2$ (1 ml $CaCl_2$ /10 ml original culture volume) and incubated on ice for 20 min. The cells were then aliquoted for use in immediate transformation experiments, or prepared for glycerol stock by a second similar centrifugation step, followed by resuspension in sterile, ice-cold 50 mM $CaCl_2$, containing 15 % (v/v) glycerol. Cells were aliquoted in chilled 1.5 ml tubes, snap-frozen in dry ice / methanol and stored at -80 °C.

2.3.3 Polymerase chain reaction

PCR reactions (a final volume of 50 µl) contained template DNA (100 ng), forward and reverse primers (100 pmol each), Vent DNA polymerase (2U) and 0.4 mM of each deoxynucleoside triphosphate (dATP, dCTP, dGTP, dTTP) in PCR buffer supplied by the manufacturer. Controls involved reactions that contained only a single primer each.

Reactions were carried out in an Eppendorf Mastercycler (Eppendorf, Germany). A standard cycle comprised:

	Hot start denaturation	(96 °C)	3 min
x30 cycles	Denaturation	(96 °C)	1 min30 s
	Annealing	(X °C)	1 min 15 s
	Elongation	(72 °C)	Y min

X is the temperature ~ 5 °C below the melting temperature of the primers. Y is the elongation time, when elongation occurs at 1 kb/min.

2.3.4 Agarose gel electrophoresis

DNA was visualised by electrophoresis in 0.8 – 1 % (w/v) agarose gel supplemented with 0.5 µg/ml ethidium bromide, made by heat-dissolving in TAE buffer (40 mM Tris acetate, 1 mM EDTA [pH 8.0]). DNA samples were prepared for electrophoresis by the addition of the appropriate amount of 6x loading buffer [40 % (w/v) sucrose, 0.25 % (w/v) xylene cyanol, 0.25 % (w/v) bromophenol blue]. DNA was electrophoresed in TAE buffer at a constant 60-100 V. Gels were run for an amount of time that facilitated adequate band separation; progress of DNA through the gel was tracked by following the xylene cyanol dye front. Ethidium bromide-bound DNA bands were visualised under a UV transilluminator. Size of fragments was estimated by calibration against a 1 kb DNA molecular mass marker run simultaneously on the gel.

2.3.5 DNA purification

2.3.5.1 Gel purification

DNA bands were excised from agarose gels and then purified using the QIAEX II Gel Extraction Kit, following the manufacturer's protocol (1999). DNA was eluted in 20 µl sterile distilled water and stored at -20 °C. The concentration of eluted DNA was estimated by electrophoresis of a defined amount of sample in an agarose gel (as in Section 2.3.4) and comparison with a DNA mass ladder, run in the gel simultaneously.

2.3.5.2 Plasmid DNA purification

Cells were cultured overnight at 37 °C with 200 rpm shaking, in 10 ml LB medium supplemented with antibiotic as appropriate. Cells were harvested by centrifugation at 6000 × *g*, 10 min, and then plasmids extracted using the Clontech (CA, USA) Nucleospin® Plus Miniprep Kit following the manufacturer's protocol (1998). Plasmids could then be visualised by electrophoresis in an 0.8 % agarose gel (Section 2.3.4)

2.3.6 A-tailing blunt-ended PCR fragments

Gel-purified PCR fragments generated by Vent DNA polymerase were A-tailed using *Taq* DNA polymerase, following the Promega A-tailing procedure protocol (pGEM-T Vector Systems Technical Manual, 1999). The reaction was carried out in 0.5 ml PCR tubes (which have low DNA binding capacity) at 70 °C for 30 min.

2.3.7 DNA ligation

Target genes were ligated into plasmid vectors using the Promega Rapid Ligation Kit and the accompanying protocol, in a total volume of 10 µl. Reactions were carried out in 0.5-ml PCR tubes and Insert:vector molar ratios always calculated, according to this protocol.

Reaction mixtures were incubated at 4 °C overnight. Ligation controls were as follows:

- Negative Ligation Control – where insert DNA was absent from the reaction mixture.
- Positive Ligation Control – where control insert DNA (supplied with the kit) was added to the reaction mixture.

2.3.8 Restriction digestion

Typical restriction digests contained 0.2 – 2 µg of plasmid DNA and were performed in 1.5 ml tubes in total volumes that ranged from 20 – 100 µl. For double digests, a buffer that was compatible for both restriction enzymes was used (details of compatibility were obtained from the manufacturer's catalogue). BSA (final concentration 100 µg/ml) was added only when recommended for an enzyme. The number of units of enzymes used depended upon the amount of DNA being cut, and was usually 10 – 20 U. Glycerol concentrations were kept below 10 % (v/v) (elevated levels of glycerol may promote star activity of some enzymes, where sequences bearing similarity to the restriction sequence are cleaved). Reactions were incubated for 1 h – overnight at 37 °C (depending on cutting efficiency of the restriction enzymes used). After incubation, samples were either stored at -20 °C, or run on an agarose gel. Target DNA fragments were excised from the gel and purified using the QIAEX II gel extraction kit (as in Section 2.3.5.1). 'Single restriction digestion' controls involved reactions containing only a single restriction enzyme per tube.

2.3.9 Transformation by heat-shock

2.3.9.1 Escherichia coii JM109 cells

Competent *E. coli* JM109 cells stored at -80 °C were thawed on ice for 5-10 min. Around 50 ng (1-5 µl) of ligation reaction (Section 2.3.7), or 1 ng plasmid DNA when vectors were being amplified, was added to 50 µl of cells, and the mixture was stirred gently with a pipette tip. Cells were incubated on ice for 30 min, transferred to a heat source at 42 °C for 1.5 – 2 min, and then immediately

returned to ice for 2 min. 950 µl of room-temperature SOC medium (LB medium supplemented with 20 mM glucose, 2.5 mM KCl, 20 mM Mg^{2+} [from stock: 1 M $MgCl_2.6H_2O$, 1 M $MgSO_4.7H_2O$]) was added, and cells were incubated at 37 °C for 1 h with shaking (maximum 180 rpm). Cells were harvested by centrifugation at 6000 × *g*, 4 °C for 2 min. Most of the supernatant was poured off, leaving ~100 µl, and the pellet was gently resuspended with a pipette tip. The entire suspension was then spread onto LB agar plates, containing antibiotic as necessary. For blue/white colour-selection of transformants containing ligated pGEM-T clones, 100 µl of 100 mM IPTG and 20 µl of 50 mg/ml X-Gal were spread onto pre-prepared LB agar plates supplemented with ampicillin (100 µg/ml). After 30 min to allow absorption, cells were spread onto the plates. Incubation was at 37 °C overnight.

2.3.9.2 Escherichia coli BL21(DE3) cells

The procedure for transformation of this strain was as above, apart from the following: approximately 1 ng plasmid DNA was added to 20 µl of cells. Cells were incubated on ice for 5 min, transferred to a heating block set at 42 °C for exactly 30 s, and then immediately returned to ice for 2 min. 80 µl room-temperature SOC medium (supplied with the expression strain, see 2.3.9.1) was added to each tube, and tubes incubated at 37 °C for 1 h with shaking (200 – 250 rpm).

2.3.10 DNA sequencing

Following amplification by PCR, genes of interest were sequenced in both directions for accuracy once ligated into the pGEM-T vector, using the Perkin Elmer ABI Prism 377 Automated DNA Sequencer. Samples were prepared by mixing 20-200 ng circular plasmid DNA, 5 pmol of a single sequencing primer (either forward or reverse) and sterile distilled water to a final volume of 6 µl. DNA was elongated using Big Dye™ Terminator Cycle Sequencing Ready Reaction kit (containing a modified *Taq* DNA polymerase) and fluorescent labels, according to the manufacturer's protocol.

2.4 Protein expression and purification

Expression was induced with a final concentration of 0.4 mM IPTG for vectors with the T7 promoter and 1 mM IPTG for vectors with the T7 *lac* promoter.

2.4.1 Small-scale expression

The protocol was an adaptation of that given in the pET System Manual (9th edition). 10 ml LB plus the necessary antibiotic/s were inoculated with around 5 transformed colonies, and incubated at an experimental expression temperature (10 – 37 °C) with 180 rpm shaking until the A_{600} = 0.4 – 1. A 1-ml sample was retained. The culture was then split into two tubes, and one culture was induced with 0.4 – 1 mM IPTG. Incubation was continued as before for up to 40 h, with 1-ml samples being taken from both cultures periodically for analysis on SDS-PAGE gels. Each sample was harvested by centrifugation (6000 × *g*, 4 °C for 5 min) and the supernatant was discarded. Cell pellets were stored at -20 °C.

2.4.2 Large-scale expression

A starter culture was prepared as follows. 10 ml LB (plus the necessary antibiotic/s) were inoculated with around 5 transformed colonies, and incubated at a chosen expression temperature (10 – 37 °C) with 180 rpm shaking, until the A_{600} = ~ 0.6. The culture was then stored overnight at 4 °C. Cells were harvested by centrifugation at 6000 × *g*, 4 °C for 5 min, and resuspended in a small amount of fresh LB and antibiotic/s. The suspension was used to inoculate 500 ml of the same medium and incubated at the chosen temperature, 180-200 rpm, until an A_{600} of ~ 0.6 was reached. The culture was induced with 0.4 - 1 mM IPTG, and incubation continued for a chosen time (determined as optimal for expression by small-scale expression experiments). Cells were harvested by centrifugation (6000 × *g*, 4 °C for 15 min) and stored at -20 °C.

2.4.3 Preparation of soluble cell extract

Frozen cell pellets were resuspended in BugBuster® Protein Extraction Reagent (5 ml/g wet cells) containing Benzonase® Nuclease (1 µl/ml BugBuster) and 1 mM PMSF, and incubated on ice for 30 min with gentle agitation. Soluble cytoplasmic extract and cell debris / insoluble material were then separated by centrifugation at 16000 × *g* at 4 °C for 20 min, and retained separately at 4 °C.

2.4.4 SDS-PAGE

Gel preparation

Proteins were electrophoresed in a 10-12.5 % acrylamide resolving gel. The resolving gel was made by mixing 4-5 ml of 30 % (v/v) acrylamide, 3 ml buffer [1.5 M Tris, pH 8.9, 0.4 % (w/v) SDS] and distilled water to a final volume of 6 ml, and was polymerised by the addition of 25 µl of 10 % (w/v) ammonium persulphate and 6.3 µl TEMED. The stacking gel was made by mixing 0.45 ml of 30 % (v/v) acrylamide, 2.4 ml buffer [0.48 M Tris, pH 6.8, 0.4 % (w/v) SDS] and 1.8 ml distilled water, and was polymerised with the addition of 25 µl of 10 % (w/v) ammonium persulphate and 5 µl TEMED.

Preparation of protein samples

Samples were prepared for loading by the addition of an equal volume of 2x loading buffer [0.05 M Tris, pH 6.8, 4 % (w/v) SDS, 20 % (w/v) sucrose, 10 % (v/v) β-mercaptoethanol, 0.2 % (w/v) bromophenol blue]. Insoluble material (Section 2.4.3) was first resuspended in 20 mM Tris buffer, pH 7.0 at approximately 40 mg/ml, and then a small amount (~ 10 µl) of this was taken and combined with loading buffer. All samples were heated to 100 °C for 3 min and then vortexed vigorously. Samples were either loaded immediately onto gels, or stored at -20 °C until use.

Acetone precipitation

Typically, a single band on an SDS-PAGE gel must contain ~ 4 µg protein to be clearly visible. When the amount of protein in a sample was too small to

visualise by SDS-PAGE, the protein was concentrated by acetone precipitation. Ice-cold acetone was added to the protein sample, to 80 % of the total volume, and the sample incubated on ice for 20 min. Precipitated protein was pelleted by centrifugation at 16000 × *g* and the liquid gently pipetted off. Any remaining liquid was evaporated by vacuum-drying. Protein was then resuspended in 2X loading buffer and an equal volume of water, and prepared for SDS-PAGE as described above.

Electrophoresis

Electrophoresis buffer was 0.052 M Tris, pH 8.3, 0.1 % (w/v) SDS, 0.4 % (w/v) glycine, and protein was electrophoresed at a constant 10 mA through the stacking gel and 20 mA through the resolving gel, until the dye front reached the end of the resolving gel. The gel was stained for 15 min in Coomassie staining solution (0.25 % [w/v] Coomassie Brilliant Blue R250, 45 % [v/v] methanol, 45 % [v/v] distilled water, 1 % [v/v] acetic acid) and destained overnight in destaining solution (20 % [v/v] methanol, 10 % [v/v] acetic acid and 70 % [v/v] distilled water).

2.4.5 Protein purification

2.4.5.1 Heat precipitation

As a preliminary purification step, soluble extract containing recombinant protein (Section 2.4.3) was heated in a water bath to 65 °C for 5 min. Precipitated proteins were removed by centrifugation at 12500 × *g* and discarded. Remaining soluble extract was then exposed to further purification steps.

2.4.5.2 His-Bind Resin chromatography

His tagged proteins were purified by His-Bind Resin Chromatography, according to the protocol in the Novagen His-Bind Kits manual (2003). Soluble cell extract was prepared for loading onto the column as in Section 2.4.3.

2.4.5.3 Gel filtration

Analytical gel filtration was carried out at 25 °C on Amersham Biosciences (Chalfont St.Giles, UK) Äkta FPLC system, using a Superdex 200 10/300GL (20-ml) column. Protein standards were: β-amylase (M_r = 200000), alcohol dehydrogenase (150000), bovine serum albumin (66000), carbonic anhydrase (29000) and cytochrome *C* (12400). Please refer to the relevant chapters for details of the different buffers used for each protein.

2.4.5.4 Anion exchange chromatography

Anion exchange chromatography was carried out on the Äkta FPLC system, using a 5-ml Q-Sepharose Hi-Trap column equilibrated with 50 mM Tris-HCl buffer, pH 8.5 (Buffer A). Protein was eluted over a 0-0.8 M gradient of NaCl in Buffer A, at a flow rate of 1 ml/min over 60 min. Fractions containing protein were stored at 4 °C in the same buffer, supplemented with 1 mM PMSF.

2.4.5.5 Protein quantification

Protein concentration was determined from the A_{280} value. Molar absorption coefficients of 37820 M^{-1} cm^{-1} for E1α, 23840 M^{-1} cm^{-1} for E1β and 24870 M^{-1} cm^{-1} for E2 were predicted by analysis of the amino acid compositions using the ProtParam programme (http://expasy.ch) (Gasteiger et al., 2005).

2.4.6 Dialysis

Dialysis tubing was prepared by two periods of boiling for 10 min, first in a large volume of 2 % (w/v) sodium bicarbonate, 1 mM EDTA (pH 8.0) and then in the same EDTA alone, with thorough rising with distilled water between and after treatments. Samples were sealed into dialysis tubing, and tubing was submerged in a large volume (usually 1 litre) of appropriate storage buffer (determined empirically) and dialysed overnight at room temperature with gentle stirring.

2.5 Enzyme assays

All materials were kept on ice, unless otherwise stated. All assays were performed in a Cary 300 Bio spectrophotometer. Kinetic parameters were determined by the direct linear method of Eisenthal and Cornish-Bowden (1974) using the Enzpack programme (Biosoft).

2.5.1 E1 assay

The assay was performed at 55 °C in 1-ml cuvettes. Buffer contained 20 mM potassium phosphate (pH 7.0), 2 mM $MgCl_2$ and 0.2 mM TPP. Buffer and recombinant E1 enzyme were pre-incubated at 55 °C for 10 min, to allow E1-TPP binding. E1 activity was monitored by following the reduction of 2,6-dichlorophenolindophenol (DCPIP; 50 µM) at 595 nm. The assay was started by the addition of up to 2 mM 2-oxoacid substrate (either pyruvate, 2-oxoglutarate, 4-methyl-2-oxopentanoate, 3-methyl-2-oxopentanoate or 2-oxoisovalerate).

2.5.2 E3 assay

The assay was performed at 55 °C in 1 ml cuvettes. Assay buffer was 50 mM EPPS buffer (pH 8.0 at 55 °C) and contained 0.4 mM of each of the substrates dihydrolipoamide and NAD^+. The reaction was started with the addition of E3 enzyme, and activity was monitored by measuring the production of NADH at 340 nm.

2.5.3 OADHC assay

The OADHC assay was performed at 55 °C in 1 ml cuvettes. Assay buffer was 50 mM potassium phosphate (pH 7.0), 2.5 mM NAD^+, 1 mM $MgCl_2$, 0.2 mM TPP, 0.13 mM CoA, and 2.6 mM cysteine/HCl. Buffer and complex (or up to 200 µl soluble cell extract) were pre-incubated at 55 °C for 10 min. The assay was started by the addition up to 2 mM 2-oxoacid substrate (pyruvate, 2-oxoglutarate, 4-methyl-2-oxopentanoate, 3-methyl-2-oxopentanoate or 3-methyl-2-oxobutanoate), and OADHC activity was monitored by measuring the

production of NADH at 340 nm. The specific catalytic activity of the reconstituted OAHDC is expressed as units per mg of E2.

CHAPTER 3: ANALYSIS OF THE PUTATIVE *THERMOPLASMA ACIDOPHILUM OADHC* OPERON, AND ITS PROTEIN PRODUCTS

3.1 Introduction

The following work describes the analysis carried out on the putative *Tp. acidophilum oadhc* gene cluster. The intention of the work was to further support standing evidence that the *Tp. acidophilum* genome contained the genes that may encode a functional OADHC, and was an important prelude to the subsequent laboratory work carried out on the enzyme complex.

3.2 The putative *oadhc* operon of *Tp. acidophilum*

The genome of *Tp. acidophilum* contains a putative operon, of 4.48 kb in length, that may encode an OADHC. The gene accession numbers are Ta 1438 (E1α), Ta 1437 (E1β), Ta 1436 (E2) and Ta 1435 (E3). However, the sequence in the database is a reverse complement, and so the gene sequences were generated using the Biology Workbench programme (http://workbench.sdsc.edu/) and then mapped onto the genome, in order to obtain the full operon sequence (See Fig. 3.1). The sequence was then examined to determine whether it contained features associated with a functional operon.

Potential of the putative *oadhc* operon to be transcribed

The transcription initiation features of archaea are more similar to those of eukaryotes than bacteria (Bell and Jackson, 2001). There is commonly an AT-rich promoter region that can sometimes stretch for many bases and it incorporates a variant of the TATA sequence that is recognized by the TATA-box binding protein (Amano et al., 2003). The TATA element is most often located around 30 bp upstream of the first open reading frame, although some

TATA elements have been identified 90 bp upstream of the ORF (Bell and Jackson, 1998, 2001; Amano et al., 2003), and transcription of the PDHC of *Corynebacterium glutamicum* is initiated 121 bp upstream of the ATG (Schreiner et al., 2005). A sequence that corresponds to the transcription factor B binding element (BRE) of eukaryotes may be found immediately upstream of the TATA element. This BRE sequence (usually two or three A bases) is important for promoter strength and is highly conserved in many archaeal promoters (Slupska et al., 2001; Bell and Jackson, 2001).

In the *oadhc* operon of *Tp. acidophilum*, there is an AT-rich region and BRE-like element that is proposed as a potential transcriptional initiation site, located 55-63 bp upstream of the first ORF. Examination of the sequence up to 150 bp upstream of the ATG start codon of ORF 1 revealed no further TATA elements. A potential poly-T stop signal is located downstream of the stop codon of the fourth ORF of the operon. No obvious TATA elements can be identified for ORFs 2, 3 or 4, internal of the operon (see Fig. 3.1).

Potential of the putative *oadhc* mRNA to be translated

Translation in archaea tends to have more in common with the mechanisms of bacteria than eukaryotes (Bell and Jackson, 2001). Three factors important in initiation of translation in archaea are the initiation codon, a sequence akin to the bacterial Shine-Dalgarno (SD) sequence and mRNA secondary structure. The most common archaeal initiation codon is AUG, although GUG and TTG are also used. An SD sequence is usually located 6-10 bp upstream of the initiation codon and may be a slight variation of GGAGG. It is complementary to, and binds, the 3' terminus of the 16S rRNA (Kawashima et al., 2000, Amano et al., 2003). However, there are some cases, for example in a number of genes studied in *Pyrobaculum aerophilum* (Slupska et al., 2001), where mRNA lacks the untranslated 5' region and SD sequence. These are termed 'leaderless transcripts'. Genes with AUG as a start codon tend to have stronger SD sequences than the GUG and TTG codons. UGA is the stop codon (Bell and

Jackson, 2001). In the *Tp. acidophilum oadhc*, a potential SD sequence (GAGG) is located 9-12 bp upstream of the AUG initiation codon of ORF 1 (E1α). Further possible SD sequences are located upstream of each ORF (see fig 3.1).

Spacing of open reading frames

The ORFs are tightly-spaced on the genome. ORFs 1 and 2, which encode the E1 α and a β polypeptides respectively, overlap and share four nucleotide bases, in an arrangement termed a 'tandem overlap' (as opposed to 'antiparallel overlap'); they reside on the same DNA strand and are read in the same direction. Overlapping genes are a common feature of microbial genomes and the feature is thought to contribute to gene conservation or the regulation of gene expression. It is suggested that the stoichiometry of the protein products of overlapping genes may be regulated through translational coupling, via site-specific programmed shifts in reading frame (Johnson and Chisholm, 2004). ORFs 3 and 4 are separated by 4 and 1 bp, respectively.

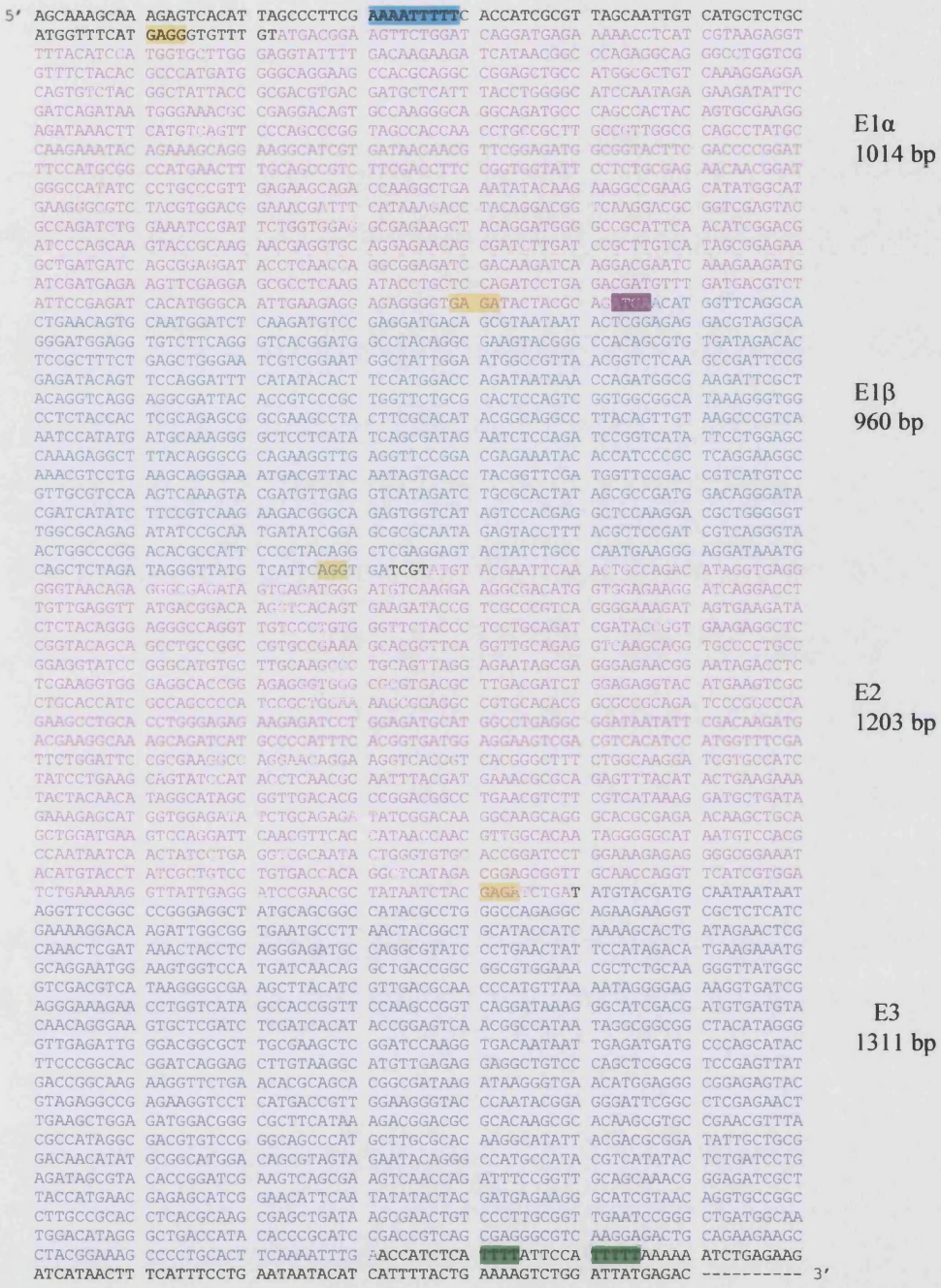


Fig. 3.1. Putative *Thermoplasma acidophilum oadhc* operon, showing genes E1α, E1β, E2 and E3. The proposed AT-rich promoter (with BRE element) and Shine-Dalgarno sequences are highlighted in blue and yellow, respectively. Two proposed poly-T stop signals are highlighted in green. Shared nucleotides of E1α and E1β are highlighted in purple. Intergene bases are shown in black.

3.3 Analysis of the E1 amino acid sequence

The amino acid sequences of E1α and E1β of *Tp. acidophilum* and a second archaeon, *Hfx. volcanii*, were aligned to those of *B. stearothermophilus* (Fig. 3.2) using the programme CLUSTALW (http://workbench.sdsc.edu/) (Thompson et al., 1994), in order to investigate the possible presence of the following proposed conserved regions. Default parameters were used in the alignments.

The TPP binding motif of E1α

Hawkins et al. (1989) describe a common sequence motif found on the E1α of 2-oxoacid dehydrogenase multienzyme complexes, and propose that this element is associated with TPP binding. The sequence is generally around 30 amino acids in length, with highly-conserved GDG and NN residues at the N- and C-termini, respectively. Internally, there is an E or D residue that immediately precedes the NN.

Conserved active-site residues in E1α

Fries et al. (2003) identify a series of conserved amino acid residues in the E1α subunit of heterotetrameric E1 enzymes that may be involved in active site functioning. The residues are situated in a region around 50 amino acids on the C-terminal side of the conserved TPP-binding motif described by Hawkins et al. (1989). They propose that the residues are involved in binding the 2-oxoacid substrate, decarboxylation of the 2-oxoacid and reductive acetylation of the lipoyl domain of E2, and phosphorylation of E1 by E1-specific kinase (in eukaryotic PDHC and BCOADHC) (Fries et al., 2003).

Conserved residues of the proton tunnel of E1α and β

The 'proton tunnel' of the heterotetrameric E1 enzyme, proposed by Frank et al., (2004), is an acidic tunnel connecting the two active sites of the enzyme. The residues in the E1 of the PDHC of *B. stearothermophilus* are αAsp180, αGlu183, βGlu28, βGlu59, βGlu88 and βAsp91 (Frank et al., 2004).

The *Tp. acidophilum* sequence

Conserved residues involved in the TPP-binding motif, active site functioning, and the proton tunnel were identified in the E1α sequence of *Tp. acidophilum* and *Hfx. volcanii* and are shown highlighted in Fig. 3.2. Conserved residues of the proton tunnel, identified in the E1β sequence of *Tp. acidophilum* and *Hfx. volcanii*, are shown highlighted in Fig. 3.3.

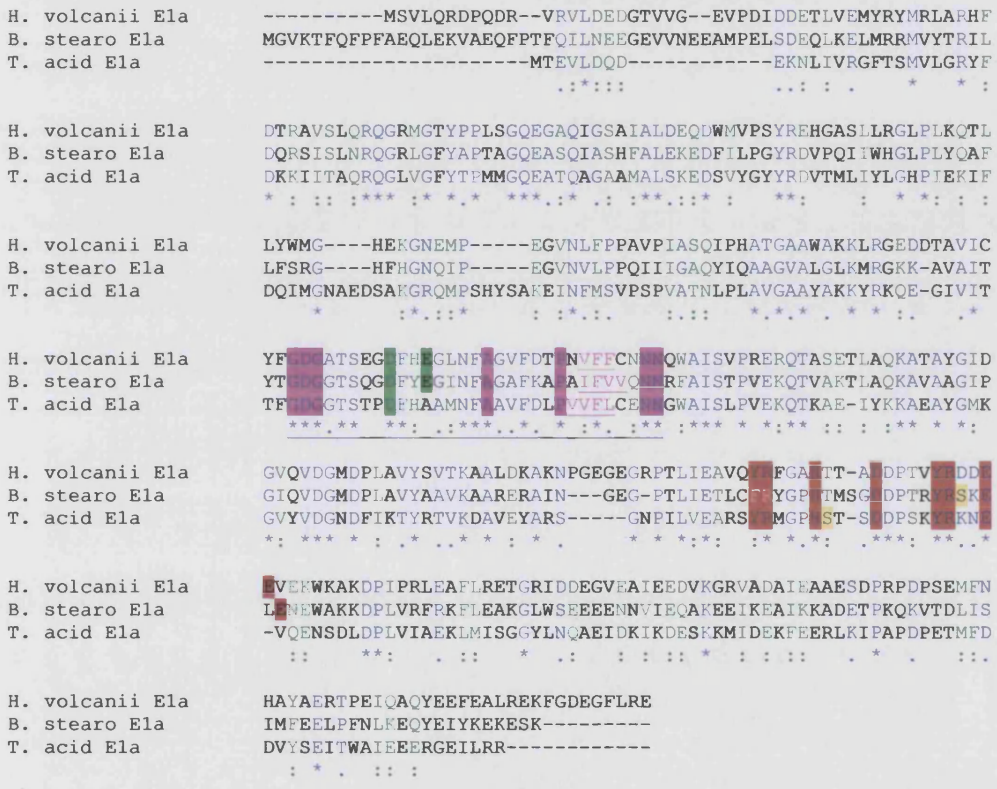


Fig 3.2. Alignment of the E1α amino acid sequences of *Hfx. volcanii, B. stearothermophilus* PDHC and *Tp. acidophilum*. Conserved residues of the proton tunnel (αAsp180 and αGlu183) are highlighted in green. Conserved residues of the TPP-binding motif are highlighted in pink, the acidic residues of the motif are underlined in pink and the entire 30 aa-region of the motif is underlined in blue. Conserved residues involved in active site functioning are highlighted in red. Two sites of phosphorylation are highlighted in yellow. With respect to the general sequence alignment, (*) denotes single, fully conserved residues, (:) denotes conservation of strong groups and (.) denotes conservation of weak groups.

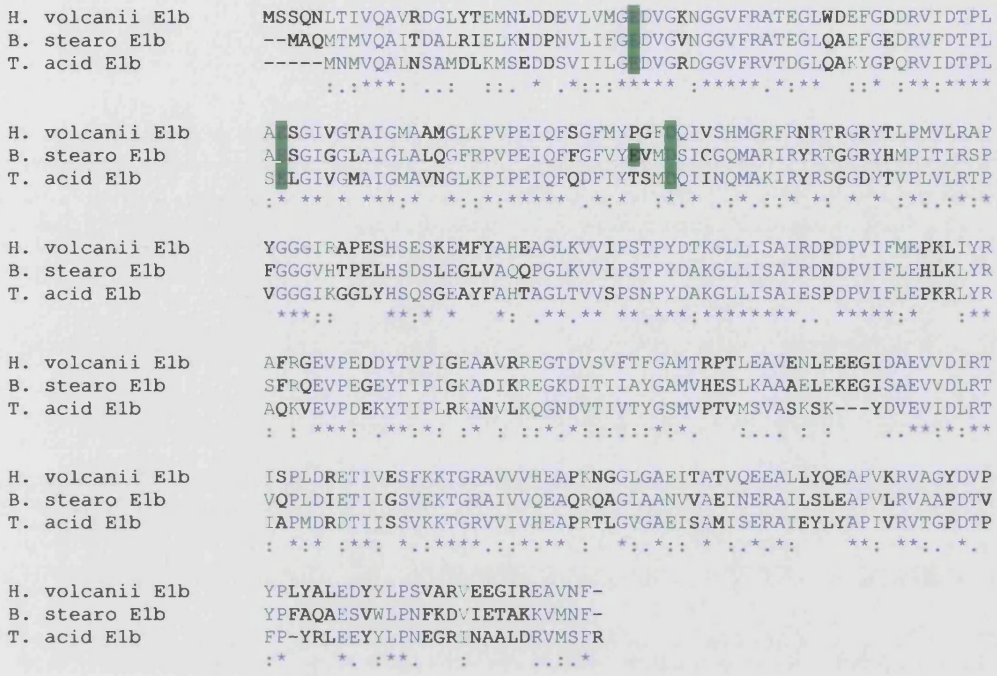


Fig 3.3. Alignment of the E1β amino acid sequences of *Hfx. volcanii*, *B. stearothermophilus* PDHC and *Tp. acidophilum*. Conserved residues of the proton tunnel (βGlu28, βGlu59, βGlu88 and βGlu91) are highlighted in green. With respect to the general sequence alignment, (*) denotes single, fully conserved residues, (:) denotes conservation of strong groups and (.) denotes conservation of weak groups.

3.4 Analysis of the E2 amino acid sequence

E2 sequence alignments

Lipoylation of a specific lysine residue in the lipoyl domain of E2 is essential for correct functioning of a multienzyme complex. The target lysine residue occupies an exposed position in a loop region of the domain (Reche and Perham, 1999). In many cases, it is flanked by moderately-conserved D and A/V residues. The amino acid sequence of the putative E2 of *Tp. acidophilum* was aligned with the E2 of the *Bacillus subtilis* BCOADHC and the *E. coli* OGDHC, using the CLUSTALW Multiple Sequence Alignment programme (http://workbench.sdsc.edu/) (Thompson et al., 1994) (Fig. 3.4). Default

parameters were used in the alignment. Through this alignment, the target lysine residue for lipoylation in the *Tp. acidophilum* E2 could be identified as lysine 42.

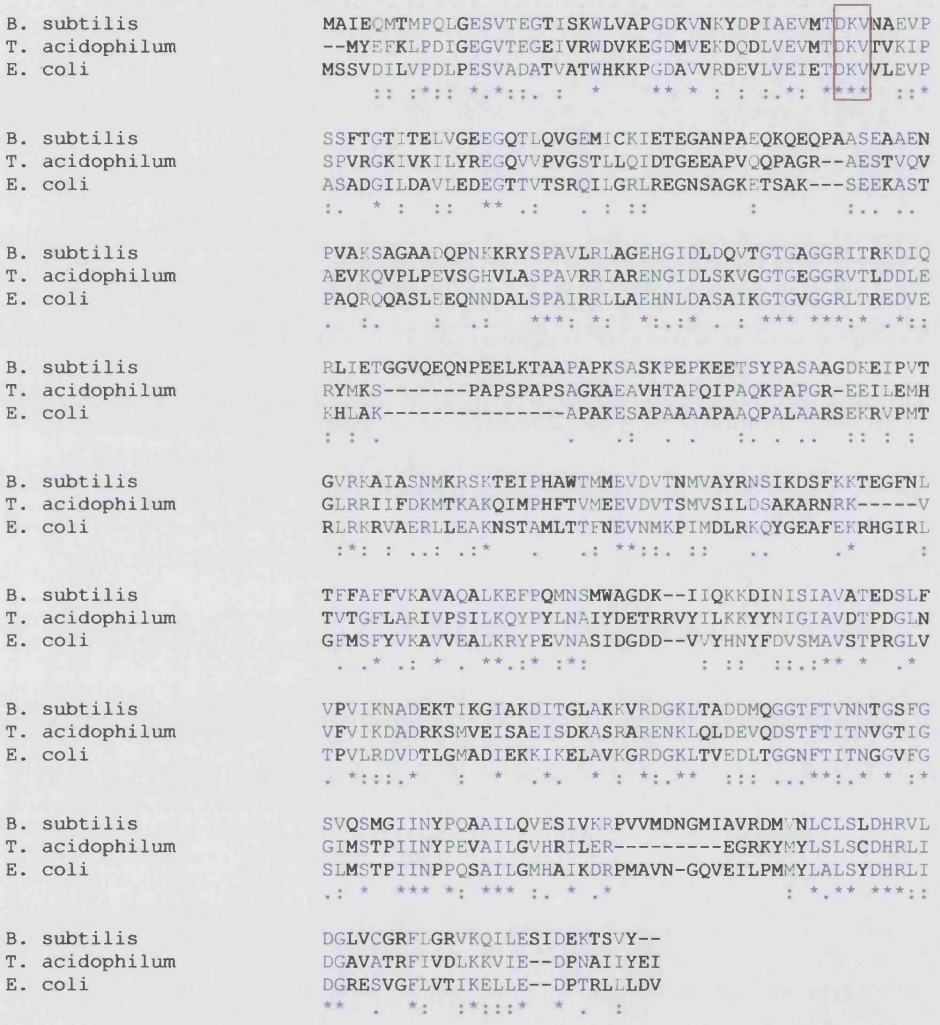


Fig. 3.4. E2 amino acid sequence alignments. *Tp. acidophilum* putative OADHC, *Bacillus subtilis* BCOADHC and *E. coli* OGDHC. The target lysine and flanking residues are highlighted in red. * indicates fully conserved regions. Semi-conserved regions are indicated (: = strong groups and . = weak groups).

Diagrammatic representation of the predicted structural domains of E2

E2 enzymes comprise between one and three N-terminal lipoyl domains (~ 80 amino acids), a peripheral subunit binding domain (PSBD, ~35 amino acids) and a C-terminal catalytic domain (~ 220 amino acids). The three domains are interconnected by proline- and alanine-rich semi-flexible linker regions, some 20-30 amino acid residues in length (Mattevi et al., 1992).

The structural domain boundaries of the *Tp. acidophilum* E2 protein that would be produced from gene Ta 1436 were predicted automatically using CD-Search (http://www.ncbi.nlm.nih.gov) (Marchler-Bauer and Bryant, 2004), and the structural prediction obtained is in accordance with the E2 enzyme of OADHCs of other organisms (Fig. 3.5; see also Fig. 1.10).

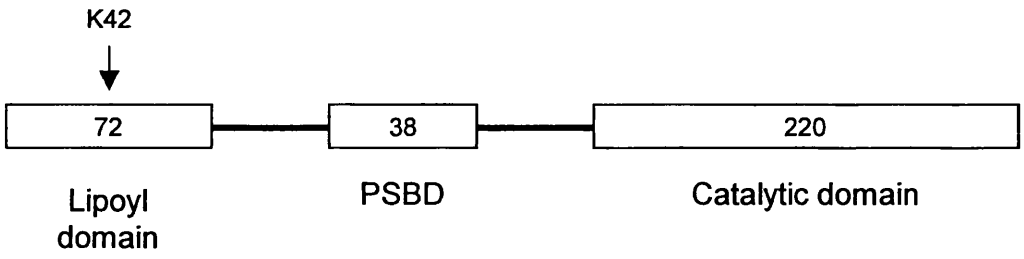


Fig. 3.5. The predicted domain structure of the E2 enzyme of *Tp. acidophilum*. Each domain is represented by an open box, with the number of amino acid residues indicated within. Linker regions, formed by semi-flexible chains of amino acid residues (black lines), connect the three domains. The lysine residue that is the proposed target for lipoylation is shown arrowed.

3.5 Putative *oadhc* operons in other archaea

The genome sequences of other members of Archaea that have recently become available on the NCBI database were examined for the presence of *oadhc* genes (and hence, Fig. 1.8 has been updated). The arrangement of the genes on each genome is shown in Fig. 3.6.

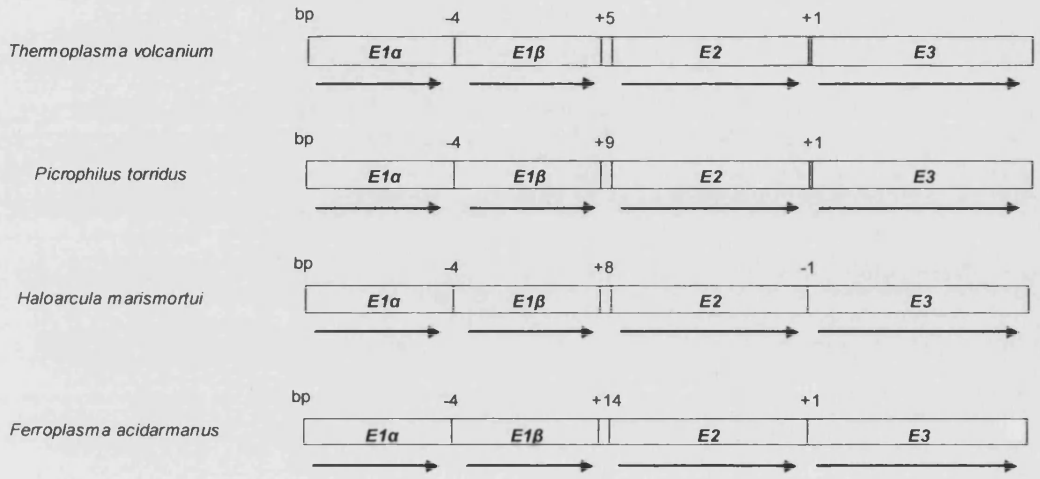


Fig 3.6. Genome analyses. The arrangement, intergene distances (bp) and proposed directions of transcription of the open reading frames (ORFs) constituting the E1α, E1β, E2 and E3 genes of the proposed 2-oxoacid dehydrogenase complex operons of *Thermoplasma volcanium, Picrophilus torridus, Haloarcula marismortui* and *Ferroplasma acidarmanus*.

3.6 Discussion

Examination of the putative *oadhc* operon of *Tp. acidophilum* provides further evidence that the cluster of genes might indeed function as a true operon. The four open reading frames encoding E1α, E1β, E2 and E3 are tightly spaced on the genome. Potential promoter regions and an SD sequence are located upstream of the first ORF, and no internal promoter sequences can be identified, indicating that the entire operon may be transcribed as a single message. Two transcriptional (poly T) stop signals are located downstream of the ORF 4. Possible SD sequences are present upstream of each ORF, indicating that each could be translated into protein. Analysis of the *Tp. acidophilum* E1 and E2 amino acid sequences, carried out both in this study and by Morgan (2001), reveal that conserved residues in their bacterial counterparts are also present in these putative enzymes.

It seems that the putative *Tp. acidophilum oacdhc* gene cluster has the capacity to encode a functional OADHC, and to be expressed *in vivo*. These hypotheses are explored in the following chapters of this study.

Further putative archaeal *oadhc* operons were identified in *Tp. volcanium*, *Picrophilus torridus*, *Haloarcula marismortui* and *Ferroplasma acidarmanus*. With the other putative archaeal operons already mentioned in Chapter 1, the *oadhc* genes have now been shown to be represented in aerobic members of both Crenarchaeota and Euryarchaeota. The putative *oadhc* genes are absent from the anaerobic archaea examined: the Thermococcales, the methanogens and *Nanoarchaeum equitans*.

CHAPTER 4: CLONING, EXPRESSION AND CHARACTERISATION OF *THERMOPLASMA ACIDOPHILUM* 2-OXOACID DECARBOXYLASE (E1)

4.1 Introduction

In the OADHC, 2-oxoacid decarboxylase (E1) carries out the first catabolic step in the overall reaction of the conversion of 2-oxoacids to their corresponding acyl-CoAs: the TPP-dependent oxidative decarboxylation of a 2-oxoacid with formation of a TPP-acyl intermediate, and the subsequent acylation of the lipoyl-lysine in the lipoyl domain of E2.

Putative E1 α- and β-encoding genes, Ta 1438 and Ta 1437, have been identified in the *Tp. acidophilum* genome. The aim of this chapter of work was to determine whether the genes do indeed encode a functional E1 protein and if so, to establish what the substrate specificity of this enzyme might be. The α and β genes were cloned and expressed using the pET vector expression system, in order to carry out a full kinetic analysis and characterisation of the recombinant protein produced.

4.2 Methods

4.2.1 E1 cloning

4.2.1.1 Cloning of E1α and E1β into pGEM-T

Primers were designed for E1α and E1β genes, such that appropriate restriction sites would be introduced into each end of the PCR-amplified gene product to allow subsequent cloning into pET vectors. The restriction enzymes *Xho* I and *Bam* HI were selected for cloning of E1α as they facilitate cloning into pET vector 19b and they do not cut the E1α gene sequence. The restriction

enzymes *Nhe* I and *Bam* HI were selected for cloning of E1β for the same reasons, to allow cloning into pET vector 28a.

Forward and reverse primers for both genes are shown below. Complementary sequences are highlighted in blue, introduced restriction sites are highlighted in red. Black sequences are "extra" bases to facilitate endonuclease attachment.

E1α:

FWD $^{5'}$CCGCTCGAGATGACGGAAGTTCTGGATCAGG$^{3'}$

REV $^{5'}$CGCGGATCCTCATCTGCGTAGTATCTCACCCC$^{3'}$

E1β:

FWD $^{5'}$CTAGCTAGCATGAACATGGTTCAGGCACTGAACAGTGC$^{3'}$

REV $^{5'}$CGCGGATCCTCACCTGAATGACATAACCCTATCTAGAGC$^{3'}$

Genes were PCR amplified from *Tp. acidophilum* genomic DNA, kindly supplied by Dr Alex Jefferies, the University of Bath. PCR amplification was carried out as described in Section 2.3.3. For E1α, the annealing temperature was 58 °C for 30 s per cycle, and elongation was 1 min 30 s per cycle. For E1β, annealing temperature was 62 °C for 30 s per cycle, elongation was 1 min 30 s per cycle. PCR products were purified by electrophoresis in a 0.8 % agarose gel, and using the Qiaex II Gel Extraction kit (as described in Sections 2.3.4 and 2.3.5).

The amplified genes were then A-tailed as described in Section 2.3.6, and separately ligated into the pGEM-T vector as in Section 2.3.7, generating the intermediate vectors pGEM-TE1α and pGEM-TE1β. Five microlitres of ligation mixture were mixed with 50 µl of competent JM109 cells, and transformation carried out as described in Section 2.3.9.1. Following incubation, five white colonies were selected as potentially containing clones. The plasmids from these colonies were extracted and visualised in a 0.8 % agarose gel, as described in Section 2.3.5.2. Plasmids that showed a size shift when compared

to circular pGEM-T control run on the same gel were treated by restriction digestion, to confirm the presence of the amplified gene insert (Section 2.3.8). For E1α, a double restriction digest was performed. Buffer was *Bam* HI buffer, and incubation was 37 °C for 2 h. A double restriction digest was also performed for E1β, buffer was Buffer E and incubation was overnight at 37 °C. Digests were electrophoresed and visualised on agarose gel. Those plasmids positive for E1α or β inserts were sequenced for accuracy as described in Section 2.3.10, harvested in greater quantities and stored at -20 °C. Larger amounts of gene insert, for ligation into pET vectors, were prepared by scaling up these restriction digestions.

4.2.1.2 Cloning of E1α into pET 19b and E1β into pET 28a

Restriction digestion of pET 19b. The vector was prepared for ligation by restriction digestion with *Xho* I and *Bam* HI. As the restriction sites for these enzymes overlap on the pET 19b vector map, it was necessary to perform two consecutive single digests. The first digest, of 20 µl total volume, comprised 17 µl (~ 1 µg) pET 19b DNA, buffer D and 10 U *Xho* I. Incubation was 37 °C for 4 h, and was followed by heat inactivation (65 °C for 20 min) and gel purification (Section 2.3.5.1). The second digest, in a final volume of 100 µl, comprised 40 µl (~ 800 ng) *Xho* I-linearised pET 19b, buffer E and 100 U *Bam* HI. Incubation was 37 °C overnight, followed by the same heat inactivation and gel purification procedures.

Restriction digestion of pET 28a. A double restriction digest was performed, in a final volume of 100 µl that contained 1 µg pET 28a DNA, Buffer E and 20 U of each restriction enzyme *Nhe* I and *Bam* HI.

Ligation into pET vectors. E1 α was ligated into pET 19b and E1 β into pET 28a using T4 DNA ligase, as described in Section 2.3.7, generating the recombinant expression vectors pET-19b-E1α and pET-28a-E1β. Again, successful ligation

was confirmed by plasmid extraction and double restriction digestion as described earlier.

4.2.2 Expression of E1α and E1β

4.2.2.1 Separate expression of E1α and E1β subunits

E. coli BL21 (DE3) pLysS cells were transformed with either plasmid pET-19b-E1α or pET-28a-E1β, as in Section 2.3.9.2. For small-scale expression of E1α, 10 ml LB supplemented with carbenicillin (50 μg/ml) and chloramphenicol (34 μg/ml) was inoculated with ~ 10 colonies from a transformation, and small-scale expression carried out as in Section 2.4.1. For E1β, the same procedure was followed, but LB was supplemented with kanamycin (30 μg/ml) and chloramphenicol (34 μg/ml). For both genes, expression was attempted at 25 ° C for up to 24 h, and 37 ° C for up to 6 h.

4.2.2.2 Co-expression of E1α and E1β subunits

Co-transformation

E. coli BL21(DE3)pLysS cells were first transformed with pET-19b-E1α, as in Section 2.3.9.2. Following overnight incubation at 37 °C, transformants were made competent for a second transformation, as described in Section 2.3.2. Media was supplemented with antibiotics (kanamycin, 30 μg/ml; carbenicillin, 50 μg/ml; chloramphenicol, 34 μg/ml). 1-4 ng of pET-28a-E1β plasmid DNA was used in the second transformation. Cells were then plated onto LB agar, containing the same antibiotics and incubated at 37 °C overnight.

Co-expression of E1α and E1β in BL21(DE3)pLysS

10 ml LB plus antibiotics (kanamycin, 30 μg/ml; carbenicillin, 50 μg/ml and chloramphenicol, 34 μg/ml) was inoculated with ~ 10 transformants, and small-scale expression carried out as in Section 2.4.1. Expression was carried out at 25 ° C for up to 24 h, and at 37 ° C for up to 6 h.

Large-scale co-expression of E1α and E1β

Large-scale expression (Section 2.4.2) of E1α and E1β, in BL21(DE3)pLysS cells, was carried out at 37 °C. Cells were induced with 1 mM IPTG at A_{600} = 0.6, and incubated for a further 6 h. Cells were pelleted and stored at –20 °C.

4.2.3 E1 purification

His-Bind Resin Chromatography and 'heat resuspension'

Soluble extract was prepared, and His tagged proteins were purified by His-Bind Resin Chromatography, as in Section 2.4.3 and 2.4.5.2. The eluted protein sample was dialysed against 1 litre of 20 mM Tris buffer (pH 9.0), containing 10 % (v/v) glycerol, with overnight stirring at 4 °C. After 24 h, precipitated protein was isolated by centrifugation (16000 × *g* for 5 min), resuspended in the same dialysis buffer and heated for 20 min at 55 °C. After this treatment remaining insoluble precipitate was removed by similar centrifugation and discarded. Purity of the E1 in the retained soluble fraction was examined by SDS-PAGE (Section 2.4.4) on a 10 % (w/v) polyacrylamide gel.

Protein concentration was determined from the A_{280} value (Section 2.4.7)

4.2.4 E1 characterisation

4.2.4.1 E1 assay

Pre-incubation experiments

To determine the length of the pre-assay incubation needed to allow E1-TPP binding, E1 was incubated in assay buffer (containing TPP) at 55 °C for 10 min, 30 min, 1 h and 4 h. After each incubation period, E1 was assayed for activity as described in Section 2.5.1.

Assay

Subsequent E1 assays involved a 10-min incubation period prior to assay.

4.2.4.2 Estimating the M_r of E1 by gel filtration

The M_r of the recombinant E1 protein was estimated by analytical gel filtration, as in Section 2.4.5.3. Buffer was 20 mM sodium phosphate (pH 7.0), 2 mM $MgCl_2$, 100 mM NaCl and 10 % (v/v) glycerol. Prior to loading of the column, 0.2 mM of TPP was added to the protein, the sample either heated at 55 °C for 5 min and then centrifuged at 16000 × *g* for 5 min, or loaded directly onto the column.

4.2.4.3 Densitometry

The E1α and E1β protein bands on SDS-PAGE gel were analysed densitometrically. The Coomassie-Blue stained gel was scanned using a Fujifilm FLA-5000 phosphorimager at 473 nm, with an LPG filter (for general-purpose excitation fluorescent digitising), and the scanned image was then analysed using the AIDA 2D Densitometry programme (Raytest, Straubenhardt, Germany).

4.2.4.4 Thermal inactivation of E1

Samples of E1 (20 µl) were incubated at 65 °C for 0, 15, 30 and 60 min. The samples from each time-point were cooled on ice, and then 10 µl used in each assay at 55 °C for activity with the substrate 3-methyl-2-oxopentanoate, as described in Section 2.5.1.

4.2.4.5 E1 temperature dependence studies

Individual E1 assays were carried out at fixed temperatures, over the range of 40 – 80 °C, as described in Section 2.5.1. A 10-min pre-incubation was carried out at the corresponding temperature, for each temperature point. The velocity of the reaction at 55 °C was taken as 100 % activity.

4.2.5 Glycerol stocks

Glycerol stocks (20 %) of co-transformed BL21(DE3)pLysS cells (containing plasmids pET-19b-E1α and pET-28a-E1β) were prepared, snap frozen and stored at -80 °C.

4.3 Results

4.3.1 E1 cloning

Following PCR amplification of the E1 genes, DNA fragments of the expected sizes of E1α (1014 bp) and E1β (960 bp) were seen (Fig. 4.1). The fragments were separately cloned into a pGEM-T vector. Restriction digestion confirmed that each gene insert had successfully ligated into its intermediate vector. The inserts were sequenced and shown to be accurate compared to the published E1α and E1β gene sequences. E1α was excised from pGEM-T and cloned into pET 19b and E1β was similarly cloned into pET 28a. Restriction digestion confirmed that both had been successfully cloned, and that the expression vectors pET-19b-E1α and pET-28a-E1β had thus been generated (Fig. 4.2).

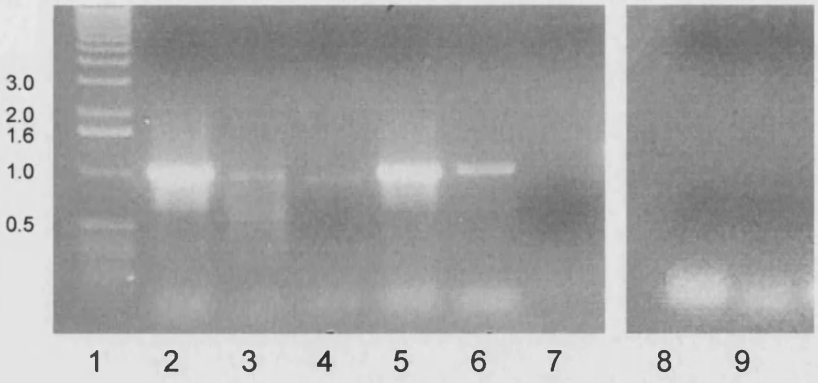


Fig. 4.1. Agarose gel electrophoresis of the PCR amplification of E1α and E1β genes. Lane 1, DNA markers (sizes given in kb); lane 2, E1α (1 kb); lane 3, E1α forward primer control; lane 4, E1α reverse primer control; lane 5, E1β (0.96 kb); lane 6, E1β forward primer control and 7, E1β reverse single primer control; lanes 8 and 9, template DNA-free controls for E1α and E1β, respectively.

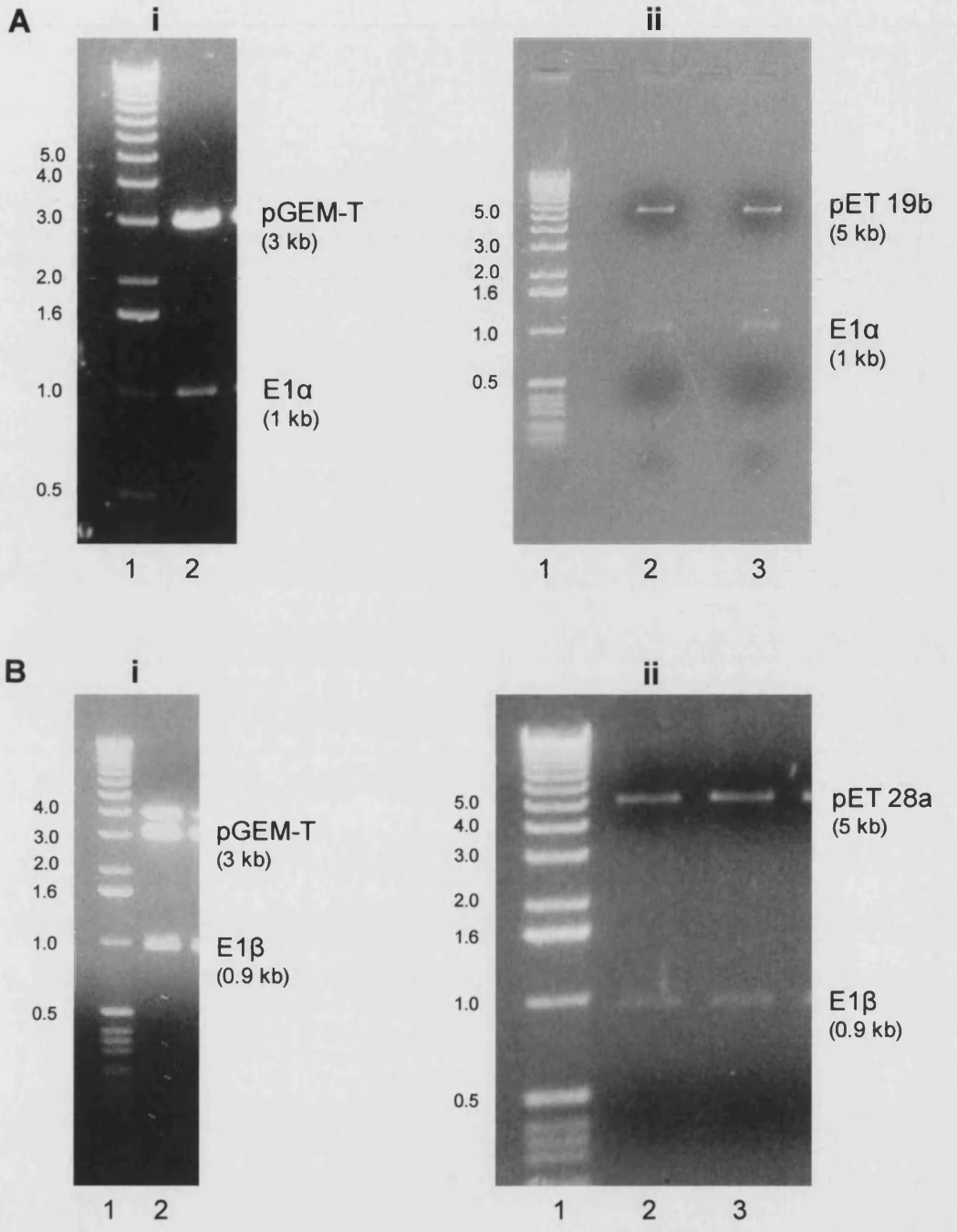


Fig. 4.2. Agarose gel electrophoresis of the restriction digestions to confirm that the gene insert is present in the vector at each stage of cloning. A: E1α cloning; B: E1β cloning. Restriction digestions of inserts from i, pGEM-T and ii, pET vectors. For all pictures: lane 1, DNA markers (sizes given in kb); lane 2/3, excised gene inserts and vectors.

4.3.2 E1 expression and purification

All expression experiments were performed at 37 °C and used the BL21(DE3)pLysS expression strain. Two different methods were used in the attempt to obtain assembled E1 protein. The first was the individual expression of E1α and E1β, to be followed by *in vitro* mixing of the purified subunits, whereas the second method involved co-expression of the genes within the same cellular compartment. In the latter, cells were co-transformed with the recombinant pET-19b-E1α and pET-28a-E1β, resulting in *E. coli* pET-19b-E1α/pET-28a-E1β co-transformants.

High levels of soluble E1α were obtained by individual expression of this subunit, but E1β expression was not detectable under the same conditions (data not shown, both soluble and insoluble fractions were examined). Co-expression experiments were, however, successful in yielding soluble E1α and E1β subunits (Fig. 4.3) and so further manipulations to the individual expression of E1β were not attempted.

Co-expressed E1αβ was purified by His-Bind Resin chromatography. Following dialysis of the His-purified E1αβ co-expressed protein into 20 mM Tris buffer (pH 9.0) containing 10 % (v/v) glycerol to remove the imidazole, the majority of protein precipitated. Resuspension of the precipitated protein at 55 °C in the dialysis buffer resulted in highly purified E1 with an approximately equal ratio of α to β polypeptides (estimated by eye). The M_r values of the α and β polypeptides were shown by SDS-PAGE to be 43000 and 37000 respectively (calculated from a plot of log M_r versus migration distances of protein standards), which compares favourably with the predicted values of 41004 and 37445 from the published gene sequences (Fig. 4.4).

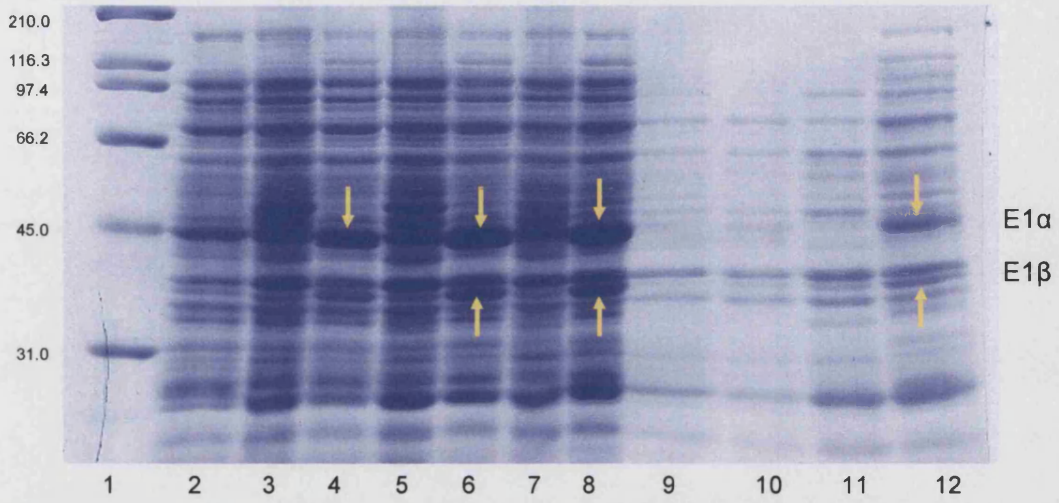


Fig. 4.3. SDS-PAGE of small-scale co-expression of E1α and β polypeptides from co-transformation. Lane 1, standard protein markers (M_r values given in kDa); lane 2, cells at the point of induction; lanes 3, 5 and 7, controls: uninduced cells at 2 h, 4 h and 6 h; lanes 4, 6 and 8, induced cells at 2 h, 4 h and 6 h; lanes 9 and 11, controls: soluble extract of uninduced cells at 2 h and 6 h. Lanes 10 and 12, soluble extract of induced cells at 2 h and 6 h. Yellow arrows indicate the progressive expression of E1 α and β, and the presence of both polypeptides in soluble form after 6 h.

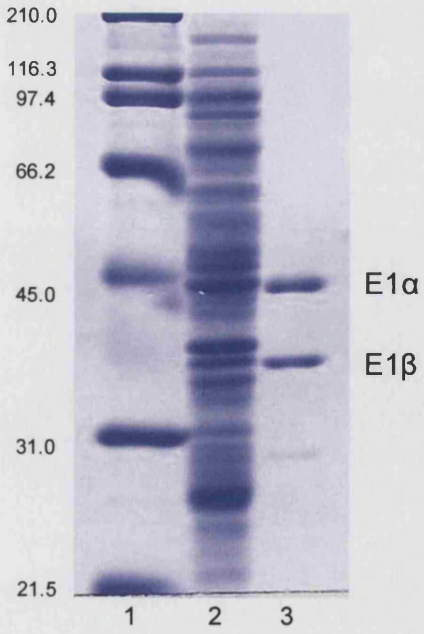


Fig. 4.4. SDS-PAGE of purified E1. Lane 1, standard protein markers (M_r values given in kDa); lane 2, soluble cell extract; lane 3, enzyme after His-Bind Resin chromatography.

4.3.3 E1 characterisation

4.3.3.1 E1 Activity

The recombinant E1 enzyme was incubated with TPP for 10 min prior to assay, a lag in the production of product being observed if the assay was initiated without prior incubation. However, increasing the pre-incubation period to 4 h resulted in no increased E1 activity over that obtained with the 10-min incubation. E1 enzymic activity was detected with the 2-oxoacids 3-methyl-2-oxopentanoate, 4-methyl-2-oxopentanoate, 3-methyl-2-oxobutanoate and pyruvate. No activity was detected with 2-oxoglutarate. For all substrates tested, the rate of reduction of DCPIP was directly proportional to the amount of enzyme in the assay, and there was a hyperbolic dependence of rate on the substrate concentration. The data were analysed using the Direct Linear method in the Enzpack programme (Biosoft), and K_m and V_{max} values determined (see figs. 4.5 – 4.12). The kinetic parameters are shown in Table 4.1.

Table 4.1. Kinetic parameters determined for purified recombinant E1 at 55 °C with branched-chain 2-oxoacids, pyruvate and 2-oxoglutarate as potential substrates. 1 Unit of enzyme activity is defined as 1 μmol DCPIP reduced per min. Values of k_{cat} are calculated per αβ dimer of the $\alpha_2\beta_2$ active enzyme. Data is based on means of three replicates.

2-Oxoacid substrate	V_{max} (U/mg protein)	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($mM^{-1}\ s^{-1}$)
3-Methyl-2-oxopentanoate	0.38 (± 0.01)	0.50	100 (± 8)	5
4-Methyl-2-oxopentanoate	0.13 (± 0.01)	0.17	42 (± 3)	4
3-Methyl-2-oxobutanoate	0.23 (± 0.01)	0.30	27 (± 3)	11
Pyruvate	0.07 (± 0.01)	0.09	596 (± 9)	0.15
2-Oxoglutarate	0	0	0	0

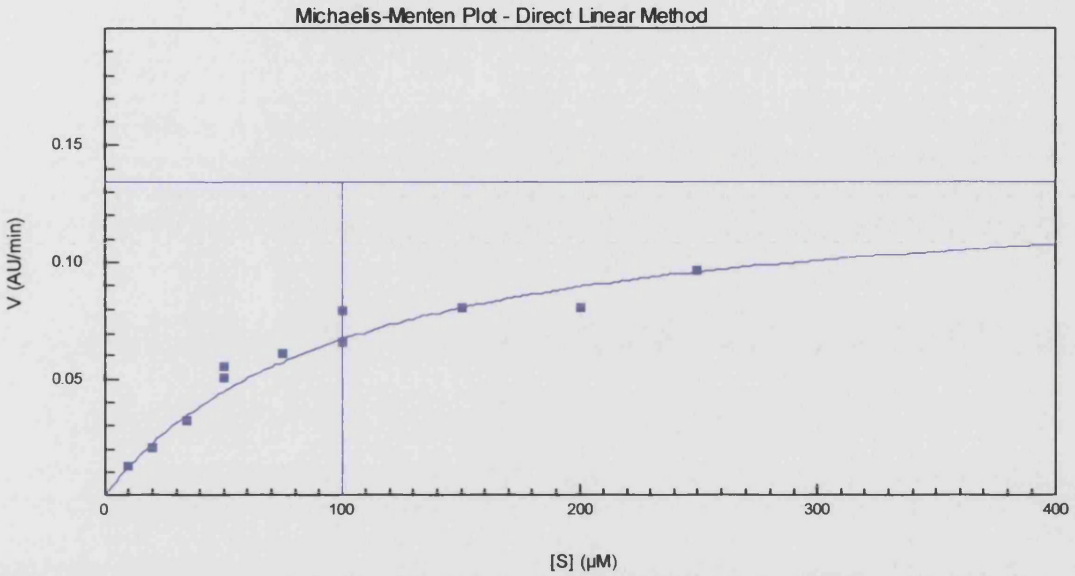


Fig. 4.5. The Michaelis-Menten plot of rate (AU/min) vs. [s] for 3-methyl-2-oxopentanoate.

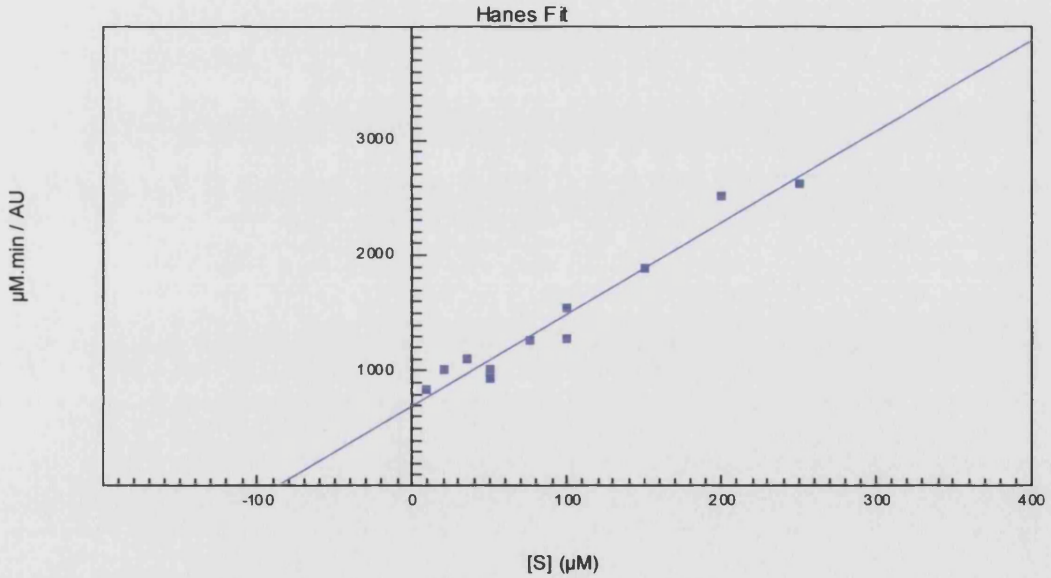


Fig. 4.6. Data obtained from E1 activity with 3-methyl-2-oxopentanoate are plotted using the Hanes Fit.

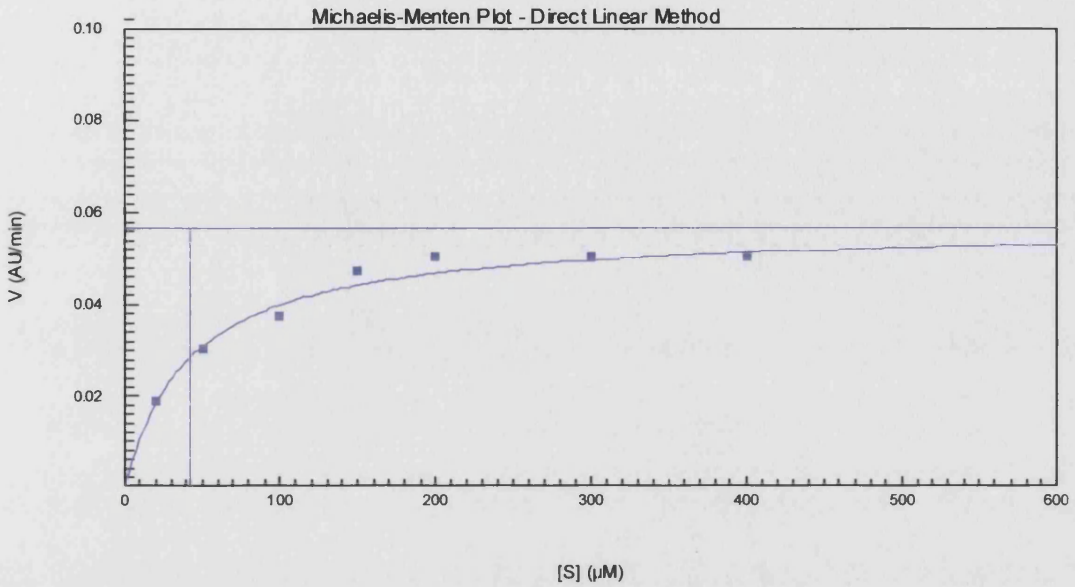


Fig. 4.7. The Michaelis-Menten plot of rate (AU/min) vs. [s] for 4-methyl-2-oxopentanoate.

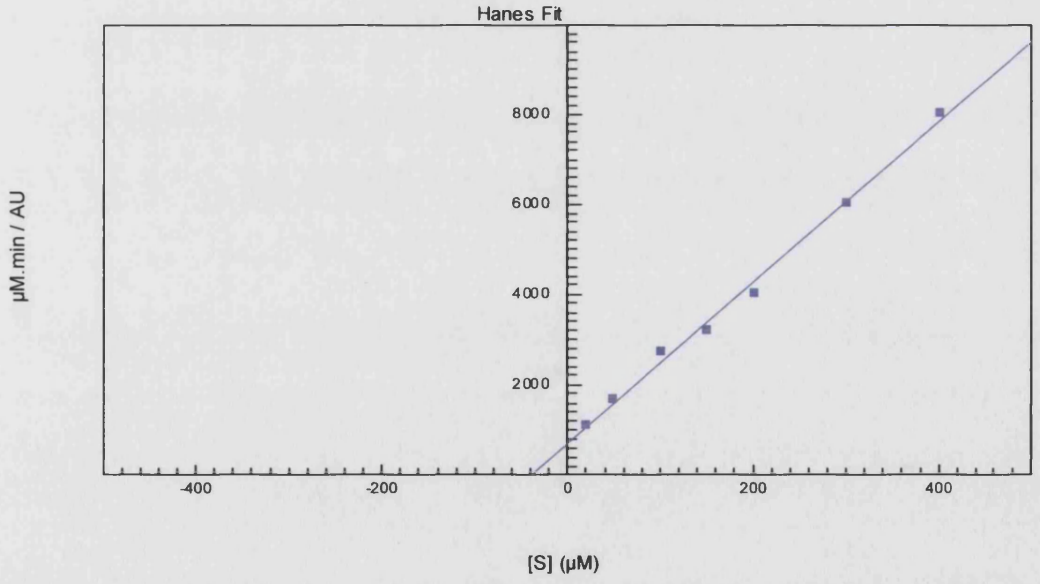


Fig. 4.8. Data obtained from E1 activity with 4-methyl-2-oxopentanoate are plotted using the Hanes Fit.

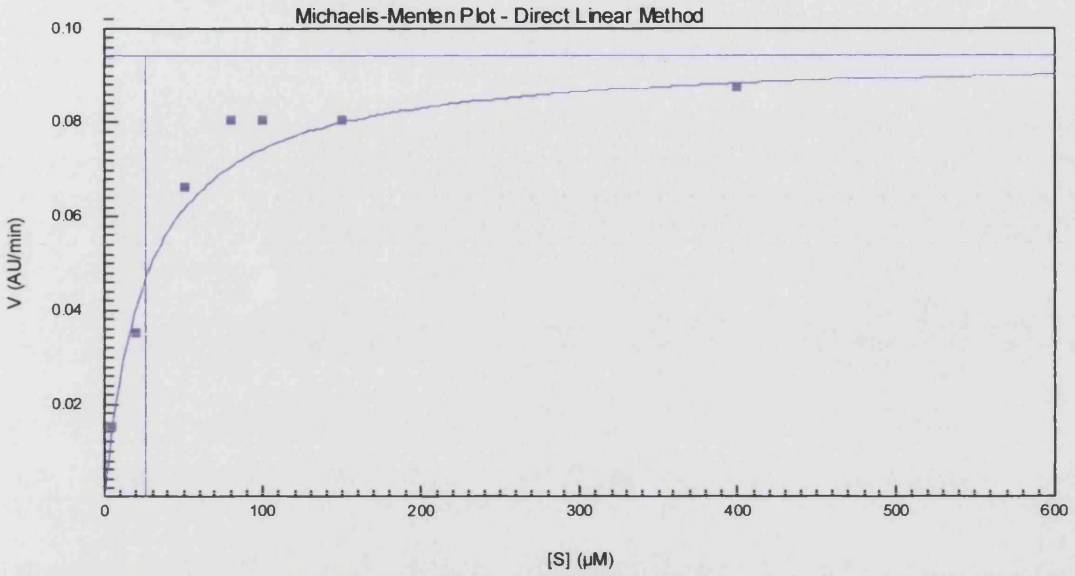


Fig. 4.9. The Michaelis-Menten plot of rate (AU/min) vs. [s] for 3-methyl-2-oxobutanoate.

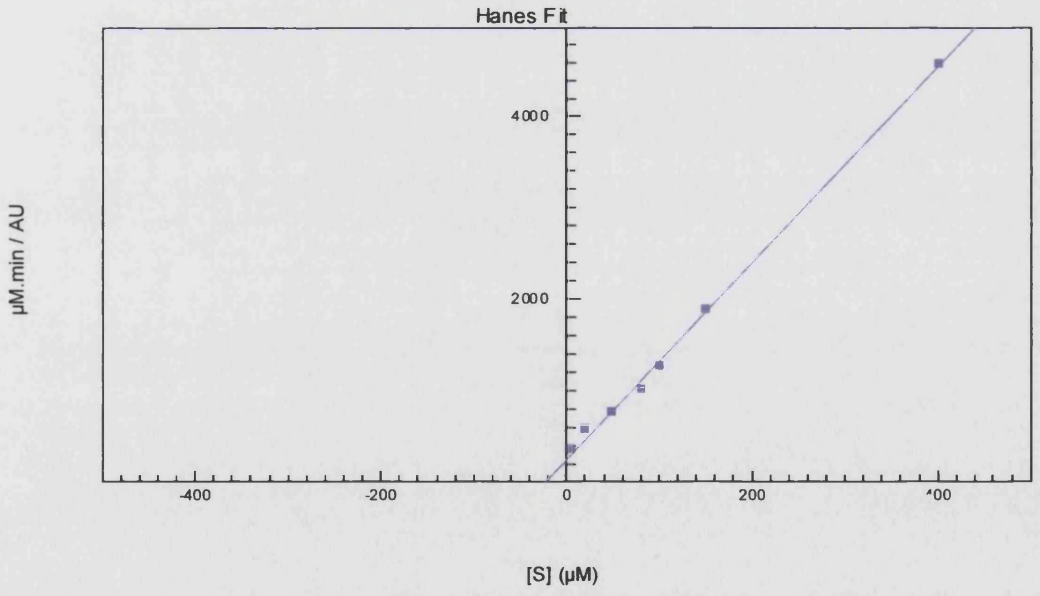


Fig. 4.10. Data obtained from E1 activity with 3-methyl-2-oxobutanoate are plotted using the Hanes Fit.

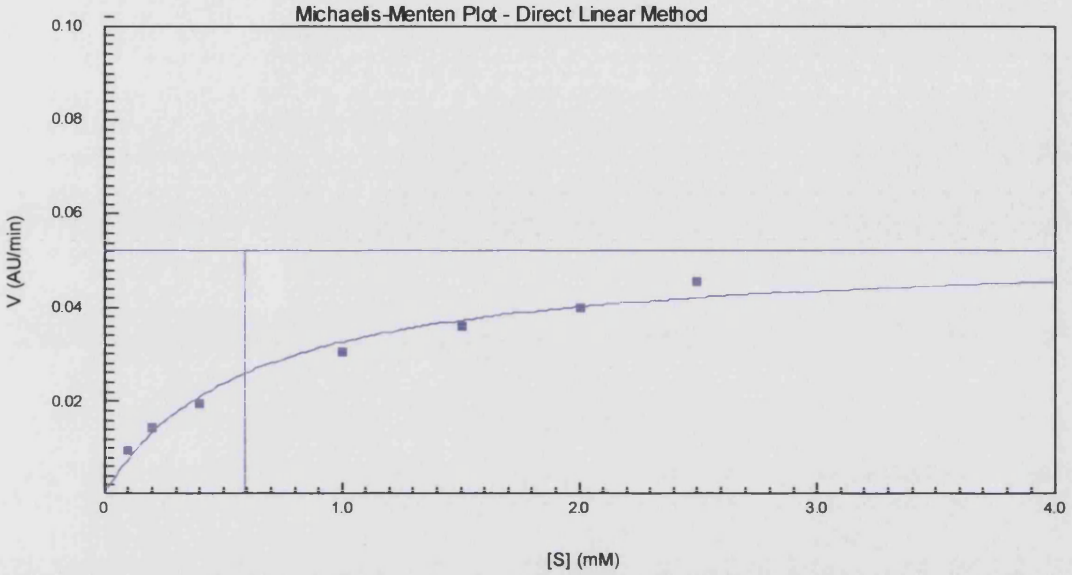


Fig. 4.11. The Michaelis-Menten plot of rate (AU/min) vs. [s] for pyruvate.

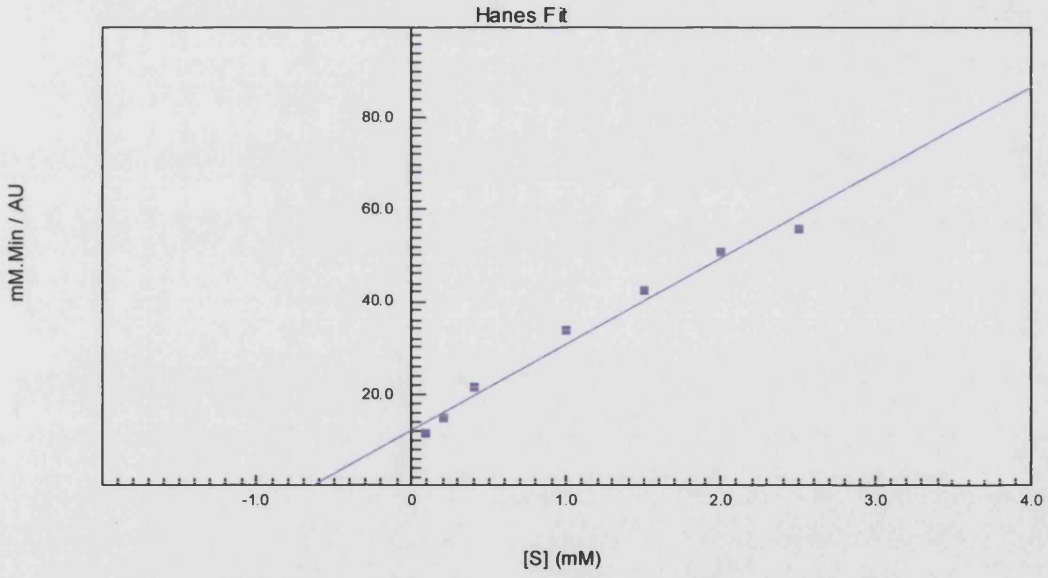


Fig. 4.12. Data obtained from E1 activity with pyruvate are plotted using the Hanes Fit.

4.3.3.2 Estimating the M_r of E1

Analytical gel filtration revealed a single peak of M_r = 165000 (calculated from the calibration curve generated from the elution of protein standards, see Section 2.4.5.3) and that was enzymically active in the E1 assay. SDS-PAGE (Section 2.4.4) of the peak fraction after gel filtration and subsequent densitometric analysis showed a 1:1 stoichiometry of the α and β polypeptides. These data suggest an $\alpha_2\beta_2$ structure for the active E1 enzyme (figs. 4.13 and 4.14). The minor peak eluting at the void volume (therefore $M_r > 1.3 \times 10^6$) was not active in the E1 assay.

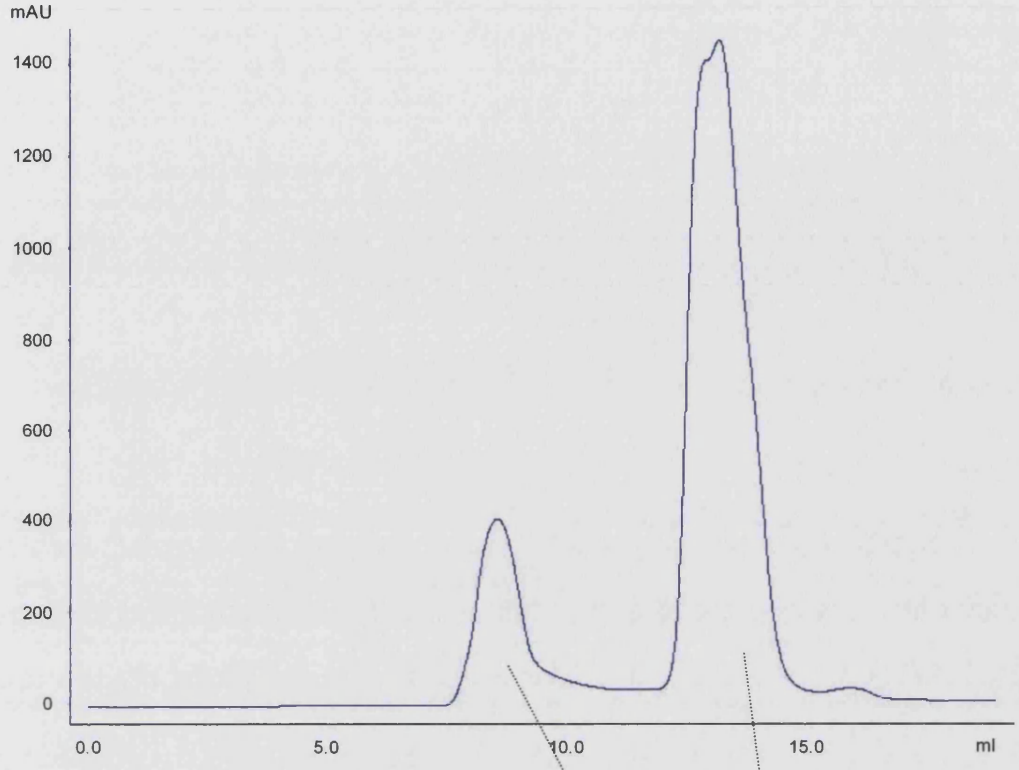


Fig. 4.13. (top): Gel filtration trace of His-purified recombinant E1 protein, showing a major and a minor peak. The major peak corresponds to an M_r of 165000 and was active in the E1 assay. The minor peak corresponds to an M_r of > 1.3 x 10^6 and was not active in the E1 assay.

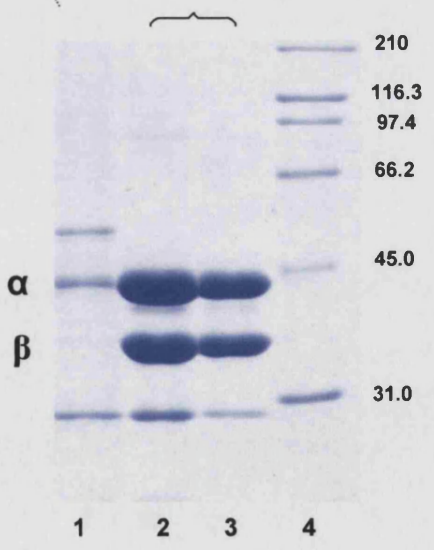


Fig. 4.14. (right): SDS-PAGE of the two gel filtration peaks. Lane 1, protein from peak 1; lanes 2 and 3, protein from peak 2; lane 4, standard protein markers (M_r values given in kDa). Densitometry was carried out on the α and β bands of lane 3. The top and bottom bands in lane 1 are protein that sometimes co-purifies with E1. The middle band is excess E1α.

4.3.3.3 Thermal Inactivation of E1

Following thermal inactivation trial experiments at 60, 65 and 70 °C (data not shown), it emerged that the most appropriate temperature for the study was 65 °C. E1 was incubated at 65 °C for 0, 15, 30 and 60 min and then assayed for activity at 55 °C. The thermal inactivation profile is shown in Fig. 4.15. The enzyme was found to have a half-life of 22 min (k = 0.021 min^{-1}) at 65 °C.

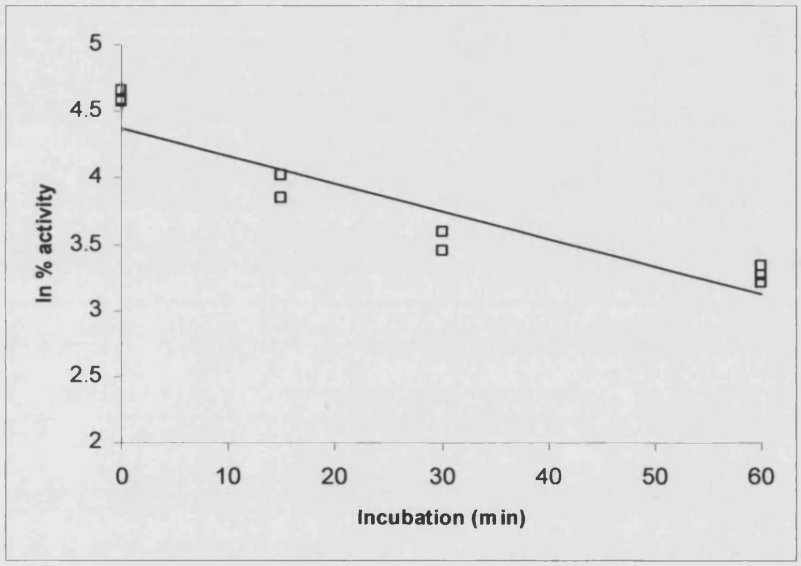


Fig. 4.15. The thermal inactivation profile of E1 at 65 °C. Thermal activity data are expressed as a percentage of that at zero time in the standard E1 assay at 55 °C.

4.3.3.4 Temperature dependence of E1 activity

Individual E1 assays were carried out at fixed temperatures, over the range of 40 – 80 °C. The dependence of E1 activity on temperature was then plotted as a percentage of the activity at 55 °C (Fig. 4.16). Maximum enzyme activity occurred between 65 and 70 °C.

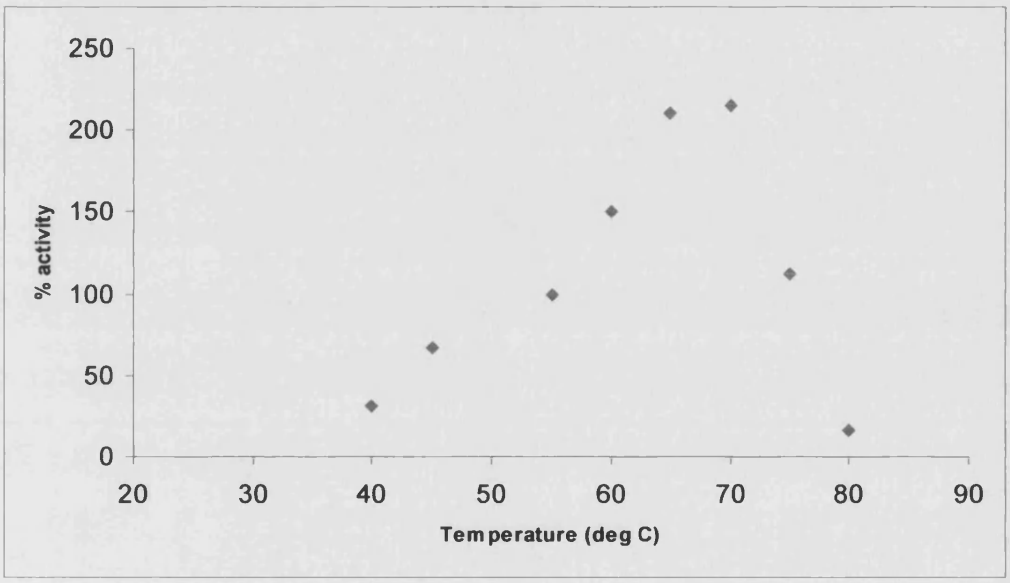


Fig. 4.16. The temperature dependence of E1 activity. Activity is expressed as a percentage of the activity at 55 °C, the standard assay temperature.

4.4 Discussion

This chapter of work demonstrates that the genes Ta 1438 and Ta 1437 in the *Tp. acidophilum* genome encode a functional enzyme that appears to be a 2-oxoacid decarboxylase (E1), the first enzyme of an OADHC. The recombinant *Tp. acidophilum* enzyme is active with the branched-chain 2-oxoacids 3-methyl-2-oxobutanoate, 4-methyl-2-oxopentanoate and 3-methyl 2-oxopentanoate, indicating that the putative complex may be a branched-chain 2-oxoacid dehydrogenase. In organisms with active BCOADHCs, these 2-oxoacids are derived from the transamination of the branched chain amino acids valine, leucine and isoleucine, respectively. However, as yet, there is no evidence for genes encoding the transaminase enzymes responsible for this conversion being present on the *Tp. acidophilum* genome.

The recombinant E1 enzyme also has activity with pyruvate, although this is 1.6-5.5 times lower than with the branched-chain 2-oxoacids (25-70 times lower when considering k_{cat}/K_m values). No activity was detected with 2-oxoglutarate. Similar activity is observed for the branched-chain 2-oxoacid dehydrogenase of bovine kidney (Pettit et al., 1978), which also has 2-5 times lower activity with pyruvate, and no activity with 2-oxoglutarate. Also, the K_m values for the *Tp. acidophilum* E1 are directly comparable with those of that complex (37-56 µM for the branched-chain 2-oxoacids, and 1000 µM for pyruvate), suggesting that the branched-chain 2-oxoacids are indeed the natural substrates for this enzyme. The k_{cat} values of the recombinant *Tp. acidophilum* E1 enzyme (0.17-0.5 s^{-1}) are similar to that reported for the recombinant E1 of *B. stearothermophilus* (0.47 s^{-1}) (Lessard et al., 1994).

Soluble E1αβ was obtained by co-transformation with the two expression vectors pET-19b-E1α and pET-28a-E1β, followed by co-expression of the α and β genes in the same cellular compartment. Co-expression has previously been shown to be successful in the production of the BCOADHC E1 of mammals (Davie et al., 1992) and *Streptomyces avermitilis* (Skinner et al., 1995), although

in those cases the α and β genes were cloned into the same expression vector. Co-expression using two separate pET vectors, in the current study, was possible because each vector conferred a different antibiotic resistance.

At the start of the chapter of work, it was also intended to attempt to generate E1 by expressing the α and β genes separately and then by mixing the polypeptides *in vitro*, a method that was successful in the production of recombinant *B. stearothermophilus* E1 (Lessard et al., 1994). However, although expression of *Tp. acidophilum* E1α yielded high levels of soluble protein, it was not possible (by SDS-PAGE analysis of both soluble and insoluble fractions) to detect any expression of E1β. Other researchers have also observed that low or zero levels of E1β protein are produced when the gene is expressed in the absence of E1α (Fisher et al., 1989; Zhang et al., 1990). It is possible that E1β may be unstable and vulnerable to degradation by proteases if it is not able to associate immediately with E1α polypeptides at the time of its expression. As the concurrent, co-expression experiments did produce soluble α and β protein, no further attempts to manipulate individual expression of E1β were made.

Gel filtration and densitometry revealed that the α and β polypeptides associate with each other, eluting together from the gel column at a point corresponding to a size of ~ 165 kDa. The equimolar ratio of α to β chains in these fractions, as revealed by densitometry, is evidence that E1 is of $\alpha_2\beta_2$ conformation. This size and conformation is characteristic of other E1 enzymes of BCOADHCs, such as that of *Pseudomonas putida* (164 kDa) and mammalian complexes (Aevarrson et al., 1999). Thermal inactivation studies revealed that recombinant *Tp. acidophilum* E1 has a half-life of 22 min at 65 °C, and maximum activity occurs at 65 – 70 °C. These values indicate that the enzyme has correctly folded upon expression, and that it demonstrates true thermophilic behaviour that might be expected of the native enzyme.

The E1α and β genes were cloned such that the polypeptide products contained N-terminal His tags. At the time of cloning, it was thought necessary to manufacture tags onto E1 to facilitate its purification from any E1 protein of the *E. coli* host cell. However, attempts to remove the tags with thrombin (pET 28a) and enterokinase (pET 19b) were unsuccessful, possibly because the cleavage sites at the N-termini of the E1 polypeptides were inaccessible to the enzymes. As the kinetic parameters of the recombinant E1 were comparable to other OADHCs, no attempts were made to reclone the gene and produce untagged E1 protein. Now that it has been determined that the recombinant *Tp. acidophilum* E1 is that of a BCOADHC, and given that *E. coli* does not possess complexes of this type, future work could involve such recloning, and purification using standard chromatographic techniques.

To conclude, when co-expressed, the Ta 1438 and Ta 1437 genes in the *Tp. acidophilum* genome encode α and β polypeptides that spontaneously assemble into a functional E1. The E1 enzyme accepts branched-chain 2-oxoacids as substrate, and is probably the first component of a functional BCOADHC. The enzyme is of $\alpha_2\beta_2$ conformation and both its size and kinetic parameters are comparable to E1 of other BCOADHCs.

CHAPTER 5: CLONING, EXPRESSION AND CHARACTERISATION OF *THERMOPLASMA ACIDOPHILUM* DIHYDROLIPOYL ACYL-TRANSFERASE (E2)

5.1 Introduction

Dihyhdrolipoyl acyl-transferase (E2) forms the structural and catalytic core of an OADHC. In bacteria and eukaryotes E2 assembles into a core molecule comprising either 24 or 60 polypeptides, depending upon the source organism and the type of complex. This chapter of work describes the cloning and expression of the putative *Tp. acidophilum* gene Ta 1436, which by sequence analysis is thought to encode an E2 protein, and the characterisation of the recombinant protein product.

5.2 Methods

5.2.1 E2 cloning

5.2.1.1 Cloning of E2 into pGEM-T

Primers were designed for the E2 gene, such that appropriate restriction sites would be introduced to each end of the PCR-amplified gene product to allow subsequent cloning into pET 28a vector. The restriction enzymes *Nde* I and *Xho* I were selected for cloning as they do not cut the E2 gene sequence.

Forward and reverse primers for both genes are shown below. Complementary sequences are highlighted in blue, and introduced restriction sites are highlighted in red. Black sequences are "extra" bases to facilitate endonuclease attachment.

FWD $^{5'}$CGCCATATGTACGAATTCAAACTGCCAGACATAGG$^{3'}$
REV $^{5'}$CCGCTCGAGTCAGATCTCGTAGATTATAGCGTTCGG$^{3'}$

The gene was PCR amplified from *Tp. acidophilum* DSM 1728 genomic DNA. PCR amplification was carried out as described in Section 2.3.3. The annealing temperature was 62 °C for 30 s per cycle, and elongation was 2 min per cycle. PCR products were purified by electrophoresis in a 0.8 % agarose gel, and using the Qiaex II Gel Extraction kit (as described in Sections 2.3.4 and 2.3.5).

The amplified DNA was then A-tailed as described in Section 2.3.6, and separately ligated into the pGEM-T vector as in Section 2.3.7, generating the intermediate vector pGEM-TE2. Five microlitres of ligation mixture were mixed with 50 µl of competent *E. coli* JM109 cells, and transformation carried out as described in Section 2.3.9.1. Following incubation, five white colonies were selected as potentially containing clones. The plasmids from these colonies were extracted and visualised in a 0.8 % agarose gel, as described in Section 2.3.5.2. Plasmids that showed a size shift when compared to circular pGEM-T control run on the same gel were treated to restriction digestion, to confirm the presence of the amplified gene insert (Section 2.3.8). A double restriction digest was performed, using Buffer D, and incubation was at 37 °C overnight. Digests were electrophoresed and visualised on an agarose gel. The plasmid positive for E2 insert was sequenced (to confirm that there were no errors introduced by PCR) as described in Section 2.3.10, harvested in greater quantity and stored at -20 °C. Large amounts of gene insert, for ligation into pET vectors, were prepared by scaling up these restriction digestions. Two consecutive single digests were performed. The first contained 5 µg pGEM-TE2 plasmid DNA, Buffer D, BSA and 30U of *Xho* I, in a final volume of 150 µl. Incubation was at 37 °C overnight, and was followed by heat inactivation (65 °C for 20 min) and gel purification (Section 2.3.5.1). The second digest, in a final volume of 200 µl, contained 1 µg *Xho* I-linearised pGEM-TE2 DNA, Buffer D, BSA and 30U of *Nde* I. DNA was gel purified as before.

5.2.1.2 Cloning of E2 into pET 28a

Restriction digestion of pET 28a

Two consecutive single restriction digests were performed. The first, in a final volume of 200 µl, contained 5 µg pET 28a DNA, Buffer D, BSA and 30 U of *Nde* I. Incubation was at 37 °C overnight, and was followed by heat inactivation (65 °C for 20 min) and gel purification (Section 2.3.5.1). The second digestion, in a final volume of 200 µl, contained 1 µg *Nde* I-linearised pET 28a, Buffer D, BSA and 30U of *Xho* I. Incubation was at 37 °C overnight, followed by the same gel purification.

Ligation into pET vectors

E2 was ligated into pET 28a using T4 DNA ligase, as described in Section 2.3.7, generating the recombinant expression vectors pET-28a-E2. Successful ligation was confirmed by the same plasmid extraction and double restriction digestion described earlier.

5.2.2 E2 expression

It proved essential to use freshly-transformed cells when expressing the E2 gene. Following all expression experiments, cells were harvested by centrifugation and stored at –20 °C.

5.2.2.1 Initial expression methods of E2

Small-scale expression trials

Small–scale expression trials were carried out at 25 °C for 24 h, or 37 °C for 5 h, 180 rpm, as described in Section 2.4.1. *E. coli* BL21(DE3), pLysS and Rosetta strains were transformed with plasmid pET-28a-E2 as described in Section 2.3.9.2. Cells were induced with 1 mM IPTG at A_{600} ~ 0.6. The medium was either LB or TB (1.08 % [w/v] tryptone, 2.16 % [w/v] yeast extract, 0.36 % [w/v] glycerol, 2.3 % [w/v] potassium dihydrogen phosphate, 12.5 % [w/v] dipotassium hydrogen phosphate), supplemented with kanamycin (30 µg/ml).

Large-scale expression

Large-scale expression was carried out as described in Section 2.4.2, in LB medium at 37 °C, 180 rpm for ~ 4 h following induction. The expression strain was pLysS.

5.2.2.2 Generating lipoylated E2

In vivo lipoylation (revised expression methods)

This method involved manipulation of expression of the E2 gene in the *E. coli* expression strain BL21(DE3). Cells were inoculated with ~ 5-10 colonies of freshly-transformed cells. The medium was LB, supplemented with kanamycin (30 µg/ml) and 0.2 mM DL-lipoic acid. No induction with IPTG took place. Cells were incubated in darkness, at 30 °C, 180 rpm for 24 h.

In vitro lipoylation, utilising *E. coli* lipoate protein ligase (lplA)

Expression of lplA

The clone carrying the *E. coli lplA* gene was kindly supplied by Professor Richard Perham of the University of Cambridge, U.K. The *lplA* gene was expressed in BL21(DE3) cells grown in LB medium supplemented with carbenicillin (50 µg/ml). Cells were cultured at 37 °C, 180 rpm for 2 h following induction at A_{600} ~ 0.6 with 0.45 mM IPTG. Simultaneously, unlipoylated E2 was expressed as in Section 6.3.

In vitro lipoylation

Soluble cell extracts were prepared from the lplA and E2-expressed cells, as in Section 2.4.3. The two extracts were mixed. In darkness, the following were added to the reaction: 1.2 mM $MgCl_2$, 1.2 mM ATP, 0.25 mM DL-lipoic acid and 1 mM PMSF. The reaction mixture was incubated at 30 °C overnight, in darkness. The extract was then subjected to His-Bind Resin chromatography, to isolate and purify the E2 protein.

In vitro lipoylation, utilising *Tp. acidophilum* lplA

Experiments to produce *in vitro*-lipoylated E2 using the recombinant *Tp. acidophilum* lplA were carried out by researchers at the Department of Biochemistry, University of Cambridge, U.K.

5.2.3 E2 purification

5.2.3.1 His-Bind Resin chromatography

Following expression of recombinant E2, a soluble cell extract was prepared as described in Section 2.4.3. For unlipoylated E2 protein (expressed in pLysS) only, the supernatant was subjected to a heat-precipitation purification step at 65 °C for 5 min, and precipitated material was removed by centrifugation at 16000 × *g* for 10 min at room temperature. The resultant soluble extract was subjected to His-Bind Resin chromatography, and fractions containing E2 protein dialysed overnight against 20 mM Tris-HCl buffer, pH 9.0, 10 % (w/v) glycerol .

5.2.3.2 Anion exchange chromatography

E2-containing fractions from His-bind chromatography were subjected to anion exchange chromatography as described in Section 2.4.5.4. Buffer A was 50 mM Tris (pH 8.5). Protein was eluted over a 0-0.8 M gradient of NaCl in Buffer A, at a flow rate of 1 ml/min over 60 min. Fractions containing E2 were stored at 4 °C in the same elution buffer, supplemented with 1 mM PMSF.

Protein concentration was determined from the A_{280} value (Section 2.4.7.)

5.2.4 E2 characterisation

5.2.4.1 Mass spectrometry to detect lipoylation of E2

E2 protein was analysed by mass spectrometry. Samples were in a 20 mM Tris (pH 9.0) buffer containing 10 % glycerol and were concentrated to ~ 15 pmol/µl in preparation for the technique.

E2 gel bands excised following SDS-PAGE were also supplied. Mass spec analysis was carried out by Dr. Catherine Botting at the University of St. Andrews, U.K.

Determination of the M_r of the E2 protein by mass spectrometry

The protein sample (10 µl, 10 pmoles/µl) was injected on to a MassPrep on-line desalting cartridge 2.1 × 10 mm (Waters), eluting with an increasing acetonitrile concentration (solutions were acetonitrile and 1 % aqueous formic acid, and an acetonitrile gradient from 2 – 98 % was applied) and delivered to an electrospray ionisation mass spectrometer (LCT, Micromass, Manchester, UK), which had previously been calibrated using myoglobin. An envelope of multiply-charged signals was obtained and deconvoluted using MaxEnt1 software to give the molecular mass of the protein.

Mapping the site of the E2 lipoylation

Protein solution (5 µl, 10 pmoles/µl) was dialysed against 50 mM ammonium bicarbonate on a VS membrane disc (Millipore), for 30 min. Sequencing grade, modified porcine trypsin (Promega) (0.3 µl, 200 ng/µl) was added and the sample incubated at 37 °C for 16 h. A portion of the sample was diluted in 5 % formic acid and the peptides separated using an UltiMate nanoLC (LC Packings, Amsterdam, Netherlands) equipped with a PepMap C18 trap & column. The eluent was sprayed into a Q-Star Pulsar XL tandem mass spectrometer (Applied Biosystems, Foster City, CA, USA) and analysed in Information Dependent Acquisition (IDA) mode. The MS/MS data generated were analysed using the Mascot search engine (Matrix Science, London, UK), with lipoylation selected as a possible lysine modification. The MS/MS spectrum corresponding to the modified peptide was also interpreted 'manually' using BioAnalyst (Applied Biosystems) tools.

5.2.4.2 Analytical ultracentrifugation of E2 complex 'core'

Lipoylated E2 protein (0.2 mg/ml), purified by His bind chromatography and anion exchange chromatography, was supplied in 50 mM Tris (pH 8.5) buffer containing 0.4 M NaCl and 1 mM PMSF. This work was carried out in collaboration with Dr. David Scott at the Centre for Macromolecular Hydrodynamics, the University of Nottingham, U.K.

The oligomeric state of E2 was studied using analytical ultracentrifugation (Hensley, 1996; Laue and Stafford, 1999). The solution heterogeneity of the protein was assessed by sedimentation velocity. In addition to the determination of the sedimentation coefficient, the data were fitted directly to the Lamm equation (Demeler, 1997) to give the M_r value of the assembled E2 protein.

5.3 Results

5.3.1 E2 cloning

Following PCR amplification of the E2 gene, a DNA fragment of the expected size of E2 (1203 bp) was seen (Fig. 5.1). The fragment was cloned into the pGEM-T vector and subsequent restriction digestion confirmed that the gene insert had successfully ligated. The insert was sequenced in pGEM-T and shown to have a nucleotide sequence identical to the published E2 gene sequence. The insert was then excised from pGEM-T and cloned into pET 28a. Again, restriction digestion confirmed that cloning had been successful, generating the expression vector pET-28a-E2 (Fig. 5.2).

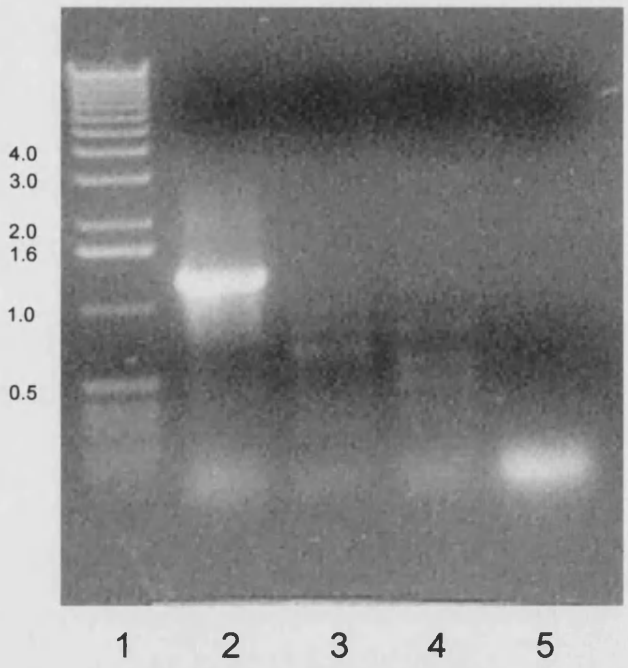


Fig. 5.1. Agarose gel electrophoresis of the PCR amplification of the E2 gene. Lane 1, DNA markers (sizes given in kb); lane 2, E2 (1203 bp); lanes 3 and 4, forward and reverse single primer controls respectively; lane 5, template DNA-free control.

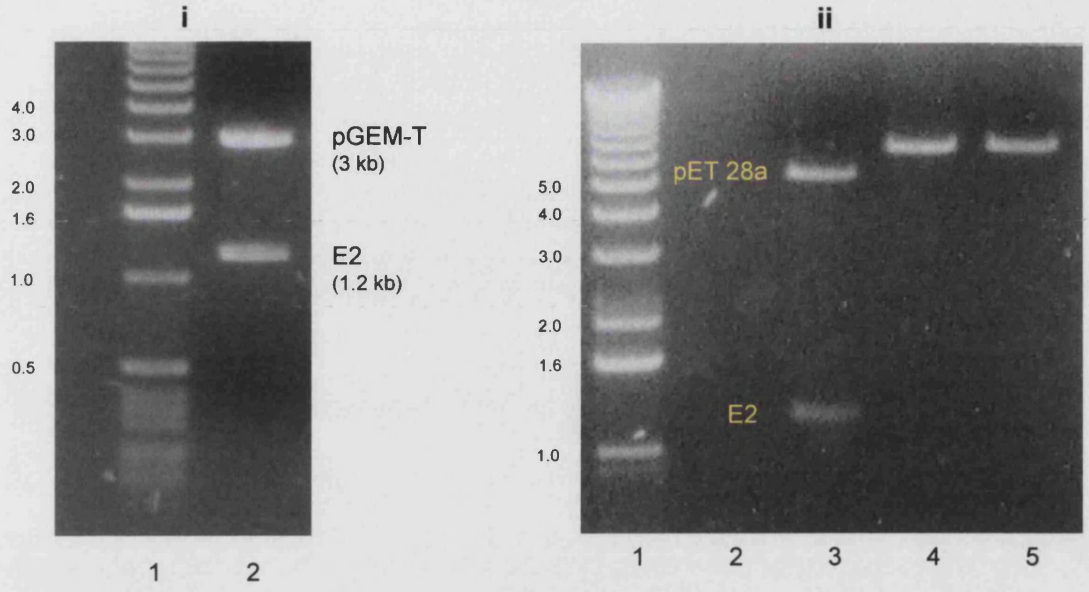


Fig. 5.2. Agarose gel electrophoresis of the restriction digestions to confirm that the gene insert is present in the vector at each stage of cloning. Restriction digestions of inserts from i, pGEM-T and ii, pET vectors. Picture i: lane 1, DNA markers (sizes given in kb); lane 2, excised gene insert and vector. Picture ii: lane 1, DNA markers; lane 3, excised gene insert and vector; lanes 4 and 5, single digest controls (*Nde* I and *Xho* I respectively).

5.3.2 E2 expression and purification

5.3.2.1 Expression and purification of unlipoylated E2 protein

Soluble E2 protein was obtained using the *E. coli* expression strain BL21(DE3)pLysS grown in LB medium, with expression taking place at 37 °C for 4 h following induction with 1 mM IPTG. The protein was purified first by heat precipitation, which removed any *E. coli* proteins with low thermostability. This was followed by His-Bind Resin chromatography and subsequently anion exchange chromatography (see Fig. 5.3). Importantly, it was noted that a large amount of E2 protein was eluted at 60 mM imidazole (corresponding with elution in the wash step). The M_r values of the E2 monomer was shown by SDS-PAGE to be 48000 (when calibrated against protein markers), which compares favourably with the predicted values of 46276 from the published gene sequence (Fig. 5.4).

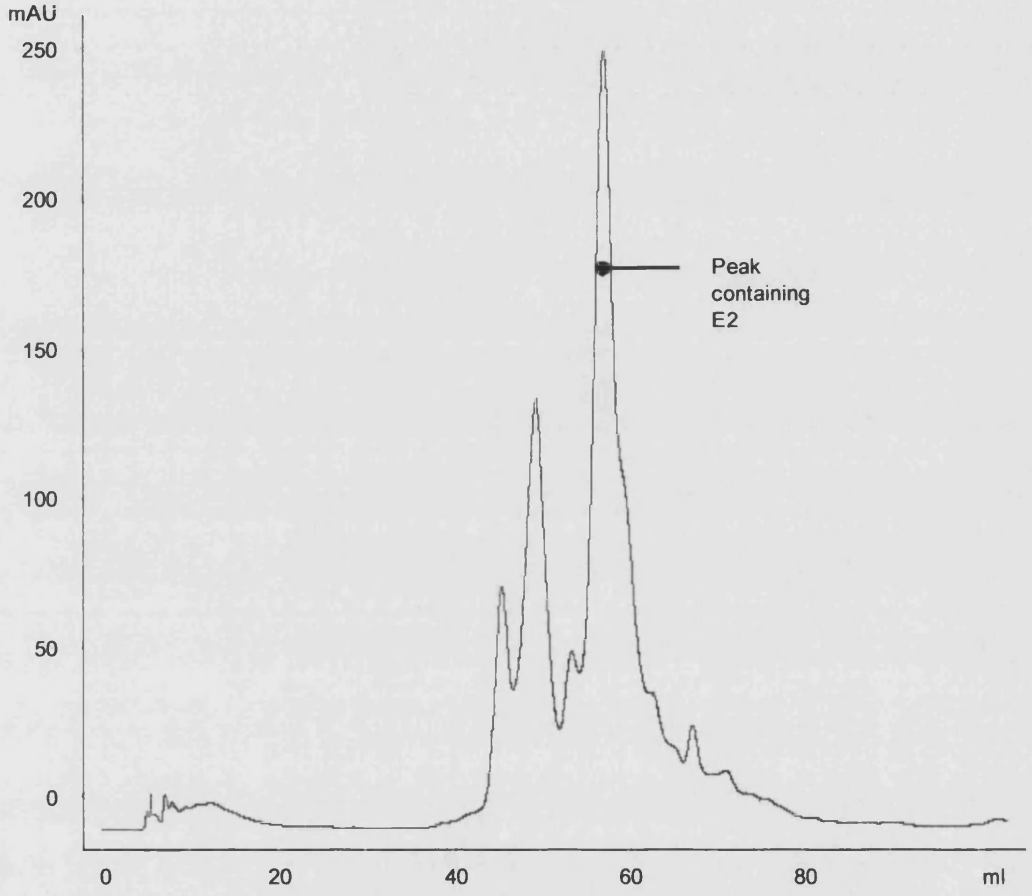


Fig. 5.3. Anion exchange chromatogram of E2 purification.

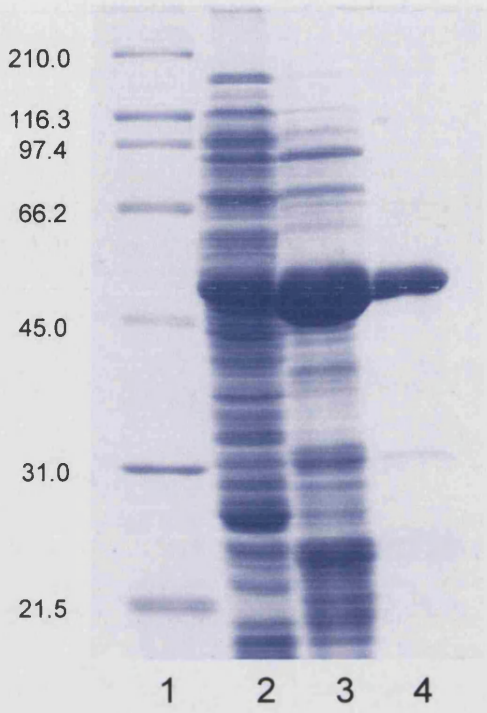


Fig. 5.4. SDS-PAGE of purified E2. Lane 1, standard protein markers (M_r values given in kDa); lane 2, soluble cell extract; lane 3, protein after His-Bind Resin chromatography (eluent from wash step); lane 4, E2 after anion exchange chromatography.

5.3.2.2 Expression and purification of lipoylated protein

In vivo lipoylation

Soluble lipoylated E2 protein was obtained by expressing the E2 gene in uninduced BL21(DE3) cells grown in LB medium supplemented with lipoic acid, at 30 °C for 20 h. No heat precipitation step was carried out, in case lipoylation of E2 had altered its thermostability. Protein was purified by His-Bind Resin chromatography and anion exchange chromatography as described earlier.

In vitro lipoylation

The overexpressed *E. coli* lplA protein was visible as an approx. 40 kDa band in soluble cell extract on SDS-PAGE gels (data not shown). However, during the *in vitro* lipoylation reaction, where soluble extracts containing E2 and *E. coli* lplA

were mixed, all the E2 protein appears to have precipitated. Attempts to lipoylate E2 *in vitro* using the *Tp. acidophilum* lplA also did not generate soluble lipoylated E2, and in those experiments E2 also precipitated during the lipoylation reaction (Edward McManus, personal communication). As *in vivo* methods were successful in yielding lipoylated E2 protein, this method was adopted for subsequent E2 expressions and there were no further attempts to improve upon *in vitro* lipoylation experiments.

5.3.3 E2 characterisation

5.3.3.1 Lipoylation analysis by mass spectrometry

The molecular weight of unlipoylated E2 protein was predicted from published gene sequences to be 46276. A 188 Da-increase in size, to 46464, would be expected when lipoylation of the protein had occurred at a single site.

E2 produced by expression without any attempt to enhance lipoylation, and also E2 produced when expression conditions were revised such to enhance lipoylation, were both analysed by mass spectrometry. For the former, results were inconclusive due to poor ionization of the sample, but no species that represented lipoylated E2 could be detected (data not shown). In contrast, the latter sample clearly contained both unlipoylated recombinant E2 protein of M_r = 46273, and lipoylated E2, M_r = 46461 (Fig. 5.5). Approximately 50 % of protein had undergone this covalent modification (Fig. 5.5). OADHC activity assays suggested that E2 is lipoylated to a greater extent when expression conditions are manipulated. Complex assembled from lipoylated E2 had a 4-fold higher specific activity than complex assembled from unmanipulated E2 (see Chapter 6 for complex assembly experiments).

Whether it was indeed the target lysine 42 residue of the lipoyl domain that had been lipoylated was analysed by trypsin digest. Fragments (and their M_r values) containing internal lysine residues as potential lipoylation targets were first predicted computationally. Trypsin digest then revealed a fragment, the size of

which would have been M_r = 1720 when unlipoylated, but whose size was M_r = 1907. The 188 Da increase of this fragment confirmed that it had undergone lipoylation, and this fragment indeed contained the target lysine-42 residue.

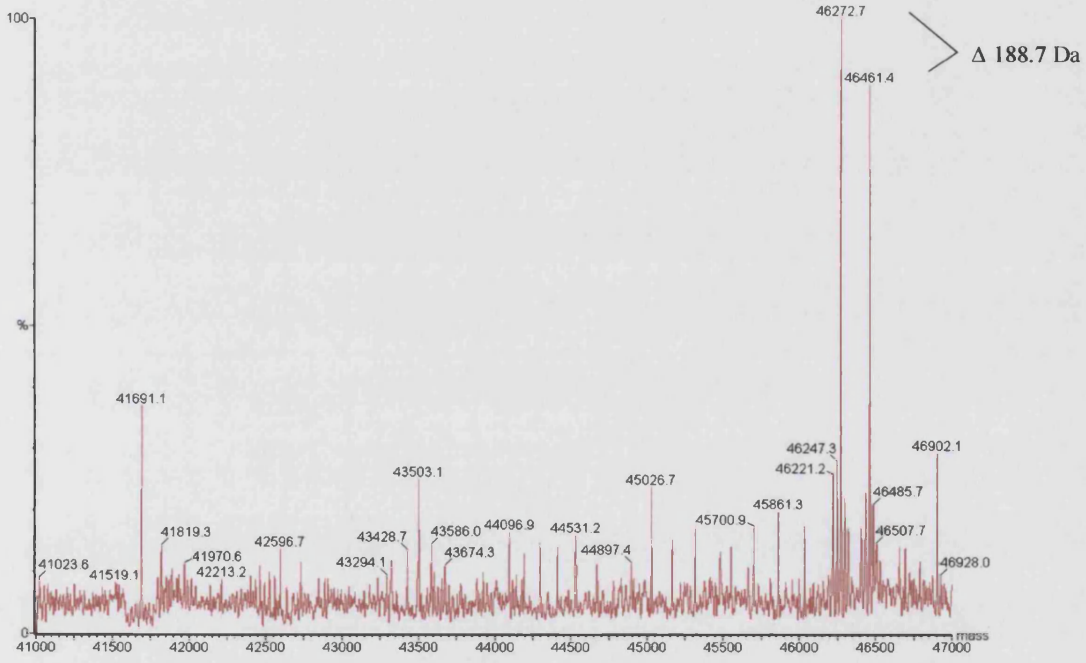


Fig. 5.5. Mass spectrometric analysis of E2 expressed with enhancement to lipoylation reveals two species, corresponding in size to unlipoylated (46 273 Da) and lipoylated (46 461 Da) E2 protein (sizes of 46276 and 46464 Da respectively, are predicted from the published amino acid sequences). Data courtesy of Dr. C. Botting, University of St. Andrews, U.K.

5.3.3.2 Estimating the size of the E2 core

E2 used in this experiment was lipoylated E2, purified by His-bind chromatography and anion exchange chromatography. Following a sedimentation velocity run, data showed a major symmetrical peak comprising 95 % of total protein, and a sedimentation coefficient of 25 S (Fig. 5.6). The M_r was determined to be 1.1 x 10^6, which from the polypeptide M_r of 46400 Da gives an oligomeric assembly of 23.7 subunits. Two minor peaks of around 18 S and 31 S were also observed. As E2 cores are known to assemble into oligomeric molecules comprising 24 or 60 copies of E2, it is reasonable to conclude from these results that the *Tp. acidophilum* core is a 24-mer.

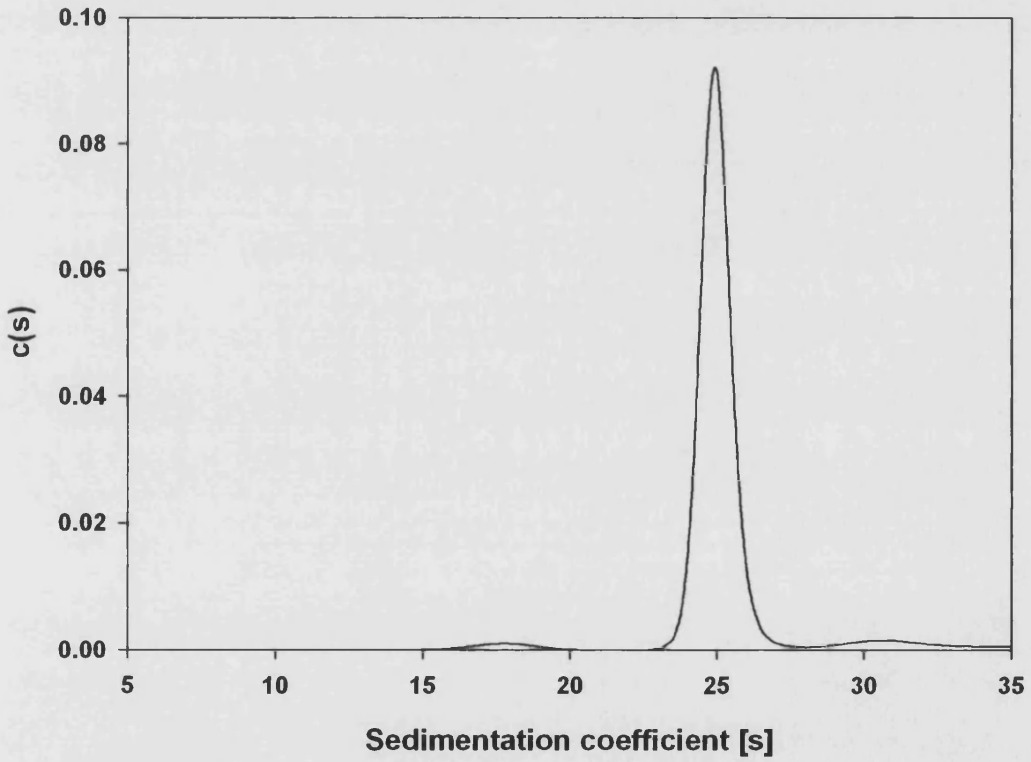


Fig. 5.6. The sedimentation velocity pattern of the E2 core, showing a major symmetrical peak at 25 S. Data courtesy of Dr. D. Scott, University of Nottingham, U.K.

5.4 Discussion

Cloning and expression

This chapter of work demonstrates that the putative gene Ta 1436 in the *Tp. acidophilum* genome encodes an enzyme that appears to be the E2 of an OADHC. The E2 polypeptides associate into a core molecule of 1.1 MDa in size, indicating that 24 E2 chains associate in formation of this core.

The E2 gene was successfully cloned and expressed as soluble protein in induced *E. coli* expression strain BL21(DE3)pLysS, but this material had not been substantially lipoylated, post-translation, by the *E. coli* lipoylation machinery. Therefore, attempts to generate lipoylated E2 were made. In *Tp. acidophilum*, the potential target lysine reside for lipoylation (lysine 42) is flanked by D and V residues, as is the corresponding lysine residue in the E2 protein of the *E. coli* 2-oxoglutarate dehydrogenase complex and of the *B. subtilis* branched-chain 2-oxoacid dehydrogenase complex.

Both *in vitro* and *in vivo* lipoylation of the recombinant E2 enzyme was attempted. In vitro methods employed recombinant lipoate protein ligase enzymes from *E. coli* and from *Tp. acidophilum*. These experiments did not prove successful; it was apparent that the majority of the E2 protein precipitated during the lipoylation reaction and further *in vitro* experiments carried out in collaboration at the University of Cambridge also failed to yield soluble lipoylated E2 protein. This does not necessarily mean that lipoylation with either of the lplA enzymes was ineffective although it has since been demonstrated that the *Tp. acidophilum* lplA is not active in the recombinant form in which it is currently produced (McManus et al., 2005). It is possible that *in vitro* lipoylation of E2 by the *E. coli* lplA was achieved during that reaction, but that the modification affected the solubility of the E2 protein and caused it to precipitate.

However, *in vivo* lipoylation of E2, by manipulation of expression conditions of *E. coli* expression strain BL21(DE3), did generate E2 protein that was around 50 %

lipoylated, as demonstrated by mass spectrometry analysis. Furthermore, the lipoylation modification had indeed occurred at the target lysine 42 amino acid residue in the predicted lipoyl domain of the protein. BL21(DE3) was selected because the control over the promoter is not as stringent as for pLysS cells, such that small amounts of recombinant protein are produced even when cells are not induced. It was hoped that the *E. coli* lipoylation machinery would be better able to cope with 'processing' this reduced amount of protein. The growth medium was supplemented with 0.2 mM lipoic acid and cells were grown in darkness. *E. coli* utilises two lipoylation pathways, exogenous and endogenous (Morris et al., 1995; Miller et al, 2000). The organism preferentially uses the exogenous pathway when lipoic acid is available in the external environment, down regulating its own internal lipoate biosynthesis (Morris, 1995 and references therein). Thus, the growth medium was supplemented with lipoic acid in the hope that this would induce the exogenous pathway of lipoylation in the expression cells. Slowing the growth temperature to 30 °C is thought to facilitate correct folding of proteins or enhance other cellular processes involved in recombinant protein expression, and would also have maintained lower expression levels of the target protein.

To further demonstrate success of the *in vivo* lipoylation technique, both lipoylated and unlipoylated E2 was assembled with E1 and E3 into complex, and the complex assayed for activity with a branched-chain 2-oxoacid as substrate (Chapter 6). The specific activity of complex containing lipoylated E2 was 4-fold higher that that containing unlipoylated E2. Taken with the mass spectrometry data, this indicates that *in vivo* lipoylation results in around 50 % lipoylated E2, and the activity of protein expressed without such manipulation corresponds to lipoylation levels of less than 12 %. The *E. coli* lipoylation machinery is probably able to lipoylate the recombinant *Tp. acidophilum* E2 to a small degree even when lipoylation is not enhanced.

Size of the E2 core

The E2 core of OADHCs is formed when E2 trimers associate into highly symmetrical molecules of cuboidal or dodecahedric shape, comprising 24 or 60 polypeptide chains respectively (Mattevi et al., 1992). Ultracentrifugation revealed an overall size for the *Tp. acidophilum* E2 core of 1.1 MDa, indicating that it must comprise an assembly of 24 polypeptide chains. It is also evident that the E2 core molecule assembles into its 24-mer state prior to the binding of E1 and E3 to the 'E1/E3 binding domain'.

To conclude, this chapter of work demonstrates that the Ta 1436 gene in the *Tp acidophilum* genome encodes an E2 protein that assembles into a 24-mer core molecule when present in homogeneous solution. The protein can successfully be lipoylated *in vivo* by the *E. coli* lipoylation enzymes.

It is reasonable to conclude that the *Tp. acidophilum* E2 is the core molecule of a BCOADHC, since the E1 enzyme encoded from the E1 genes of the same gene cluster is of the $\alpha_2\beta_2$ conformation characteristic of the E1 of BCOADHCs, and that a 24-mer core configuration is characteristic of BCOADHC E2 molecules (Heath et al., 2004 [see Chapter 5]; Aevarrson et al., 1999; Pettit et al., 1978),

CHAPTER 6: ASSEMBLY AND ASSAY OF THE 2-OXOACID DEHYDROGENASE MULTIENZYME COMPLEX

6.1 Introduction

This chapter describes the *in vitro* assembly of the putative *Thermoplasma acidophilum* OADHC from the recombinant E1, E2 and E3 subunits. The cloning, expression and characterisation of E1 and E2 were described in chapters 4 and 5. E3 was cloned, expressed and characterised in this laboratory by Hans Christian Aass (Aass, 2003), although further characterisation of E3 has been carried out in the current study and is described at the beginning of this chapter.

6.2 Methods

6.2.1 Cloning, expression and purification of E3

**This methods were established by Hans Christian Aass for the degree of MPhil (2003) and methodologies are described in detail in that thesis.*

*The E3 gene was cloned into pET 28a and expressed in BL21(DE3) cells. For expression of E3 protein, LB supplemented with kanamycin (30 µg/ml) was inoculated with cells from a glycerol stock. Cells were not induced with IPTG, and expression was for 5 h at 37 °C and 180 rpm after the A_{600} had reached a value of 0.6. E3 protein was then subjected to a 10-min, 65 °C heat-precipitation purification step, and then purified by His-Bind Resin chromatography. E3 was dialysed overnight against 20 mM Tris (pH 8.5) buffer.

6.2.2 E3 characterisation

6.2.2.1 E3 assay

The assay is described in Section 2.5.

6.2.2.2 Determination of the M_r value of E3

E3 was analysed by gel filtration, as described in Section 2.4.5.3. The buffer was 20 mM sodium phosphate (pH 7.0), 100 mM NaCl and 10 % (v/v) glycerol. E3 was incubated at 55 °C for 5 min, and then centrifuged at 12500 × *g* for 5 min prior to loading on the column.

6.2.3 OADHC assembly and assay

6.2.3.1 OADHC assay

The assay is described in Section 2.5

6.2.3.2 Stoichiometry experiments

Experiments were performed to determine the optimum molar stoichiometry of E1:E2:E3 to be combined together during *in vitro* assembly of the OADH complex. In the first study, the molar stoichiometry of E1 was varied, with E2:E3 remaining at 1:1. The stoichiometries of E1:E2:E3 were 0.5:1:1, 1:1:1, 2:1:1, 3:1:1, 4:1:1 and 6:1:1. In the second study, the molar stoichiometry of E3 was varied, with E1:E2 remaining at 3:1. The stoichiometries of E1:E2:E3 were 3:1:0.05, 3:1:0.1, 3:1:0.25, 3:1:0.5 and 3:1:1. Complex was assayed for activity using the substrate 3-methyl-2-oxopentanoate.

6.2.3.3 Pre-incubation experiments

Complex was assayed for activity using the substrate 3-methyl-2-oxopentanoate. The molar stoichiometry of E1:E2:E3 was 3:1:0.1.

<u>Incubating enzymes separately</u>

A series of experiments was performed, where complex was part-assembled by mixing two of the three component enzymes, in every possible combination, and incubating for 1 h at 55 °C. The third enzyme was then added, and the complete complex pre-incubated for 10 min at 55 °C and assayed as normal.

Varying incubation temperature and length

Complex was incubated for either 1 h at 55 °C, 1 h at room temperature or overnight at 4 °C. The complex was then pre-incubated for 10 min at 55 °C and assayed as normal.

6.2.3.4 OADHC purification by gel filtration

Complex was assembled in a 3:1:0.1 molar stoichiometry, with the addition of 0.2 mM TPP. The mixture was incubated for 5 min at 55 °C and then centrifuged, 12500 × *g* for 5 min prior to loading on the column. Gel filtration was carried out, as described in Section 2.4.5.3. Buffer was 20 mM sodium phosphate (pH 7.0), 2 mM $MgCl_2$, 100 mM NaCl and 10 % (v/v) glycerol. Peak fractions were assayed for E1, E3 and OADHC activity and subjected to SDS-PAGE (Section 2.4.4).

6.3 Results

6.3.1 Expression and purification of E3

All of the following work was carried out by the author, using the E3 clone manufactured by Hans Christian Aass (2003). E3 was expressed in uninduced BL21(DE3) cells, and the recombinant protein purified by heat-precipitation and His-Bind Resin chromatography. The predicted M_r of E3 from published gene sequences was 49867. Accordingly, SDS-PAGE analysis showed the M_r of the recombinant E3 to be 52000 (Fig. 6.1), calculated from a plot of log M_r versus migration distances of protein standards.

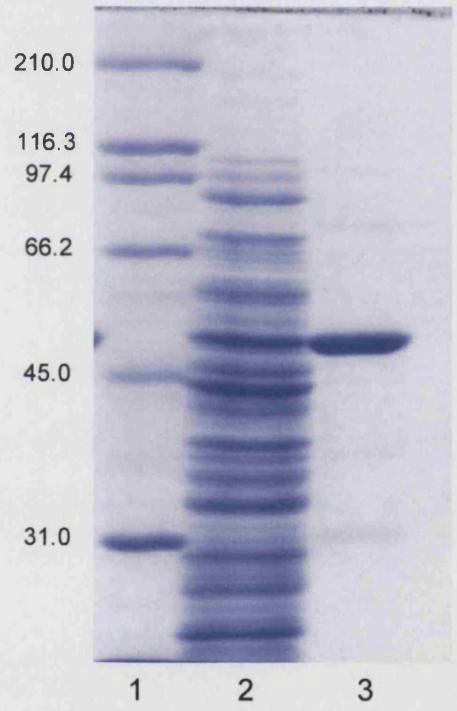


Fig. 6.1. SDS-PAGE of expression of E3. Lane 1, standard protein markers (M_r values given in kDa); lane 2, soluble cell extract; lane 3, His-bind purified E3.

6.3.2 E3 characterisation

6.3.2.1 E3 activity

E3 activity was assayed with dihydrolipoamide (DHlip) and NAD^+, and the enzyme showed a hyperbolic dependence of rate on the concentration of these

substrates. The rate of production of NADH was directly proportional to the amount of enzyme in the assay. Kinetic parameters for the enzyme were determined and are shown compared to those of *Haloferax volcanii* recombinant E3 (Jolley et al., 1996) in Table 6.1.

Table 6.1. The kinetic parameters of recombinant *Tp. acidophilum* and *Hfx. volcanii* E3.

	V_{max} (U/mg)	K_m NAD^+ (μM)	K_m DHlip (μM)
Tp.acidophilum E3	NAD^+: 22 (±0.2) [a] DHlip: 22 (±0.8) [a]	28 (±0.1)	197 (±0.1)
Hfx. volcanii E3 (Jolley et al., 1996)	24[bc]	80	18

[a] assays carried out at 55 °C
[b] assays carried out at 37 °C
[c] specific activity value, NAD^+ and DHlip at saturating concentrations

6.3.2.2 Determination of the M_r value of E3

Analytical gel filtration revealed a major peak of M_r = 99000 that was enzymically active in the E3 assay. From the published amino acid sequence, an E3 dimer would be 99.7 kDa. Two further, merged peaks corresponding with a size range of 217 – 431 kDa were also active in the E3 assay. SDS-PAGE analysis revealed the presence of E3 across all three peaks (figs. 6.2 and 6.3).

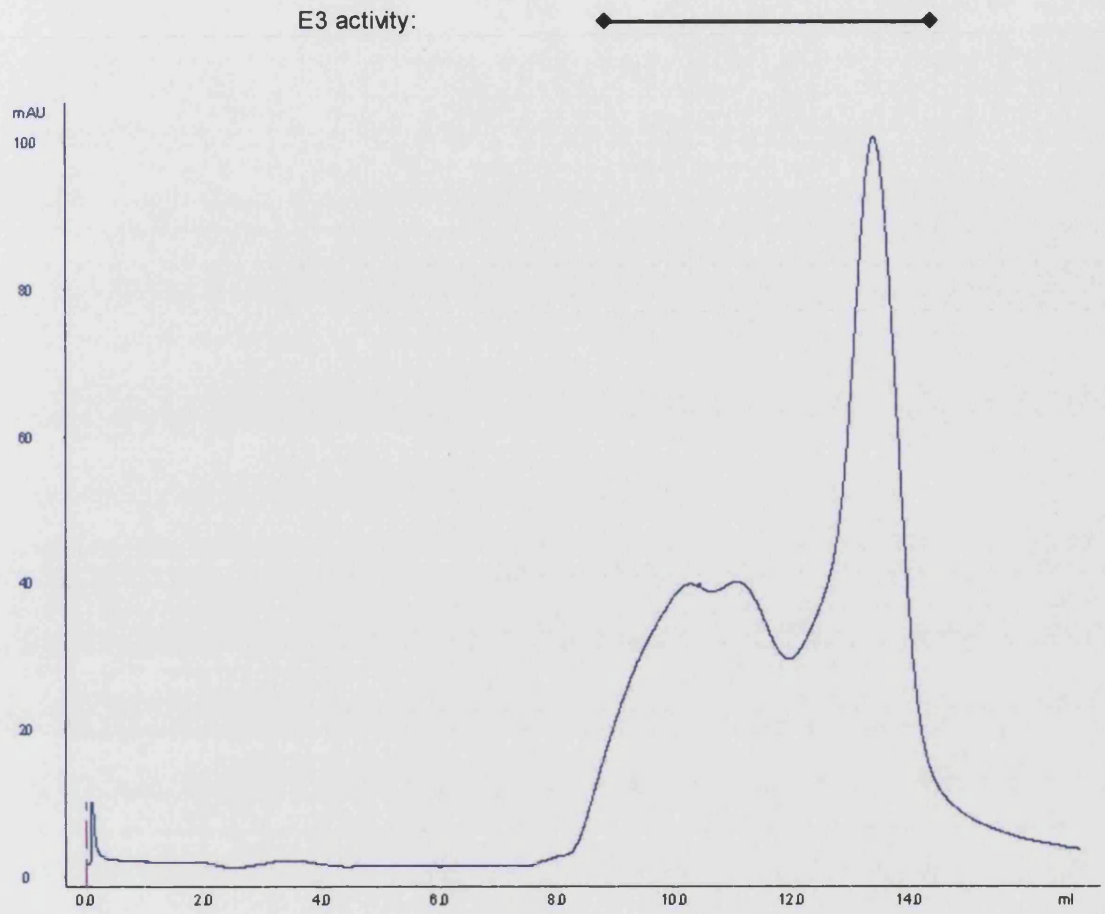


Fig. 6.2 (top). Gel filtration trace of His-purified recombinant E3 protein. A major and two minor peaks are seen. The major peak corresponds to an M_r of 99 kDa, the combined minor peaks correspond to a size range of 217 – 431 kDa. All peaks were active in the E3 assay, as indicated on the graph.

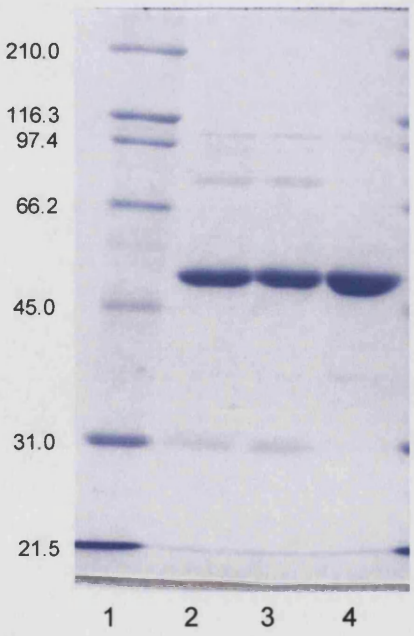


Fig. 6.3. (right). SDS-PAGE of E3 protein in gel filtration peaks. Lane 1, standard protein markers (M_r values given in kDa); lane 2, protein from minor peak 1 (eluting at 10 ml); lane 3, protein from minor peak 2 (eluting at 11.5 ml); lane 4, protein from major peak. All peak fractions were active in the E3 assay.

6.3.3 Assembly of the OADHC and complex activity

6.3.3.1 Stoichiometry experiments

These experiments were carried out to determine the optimum molar stoichiometry of E1:E2:E3 to be mixed together for *in vitro* assembly of the OADH complex.

Varying the molar stoichiometry of E1

A series of assays was performed, where the molar stoichiometry of E2 and E3 remained fixed, and the amount of E1 was varied. The stoichiometries of E1:E2:E3 were 0.5:1:1, 1:1:1, 2:1:1, 3:1:1, 4:1:1 and 6:1:1. Complex was assayed for activity against the substrate 3-methyl-2-oxopentanoate, and activity plotted as a function of the molar ratio of E1. Complex activity increased as the ratio of E1 increased from 1 to 3, but no further increase was seen at higher E1 ratios (Fig. 6.4). Assuming a line of best fit through the data points, an optimum molar ratio of E1:E2 appeared to be around 2:1-3:1.

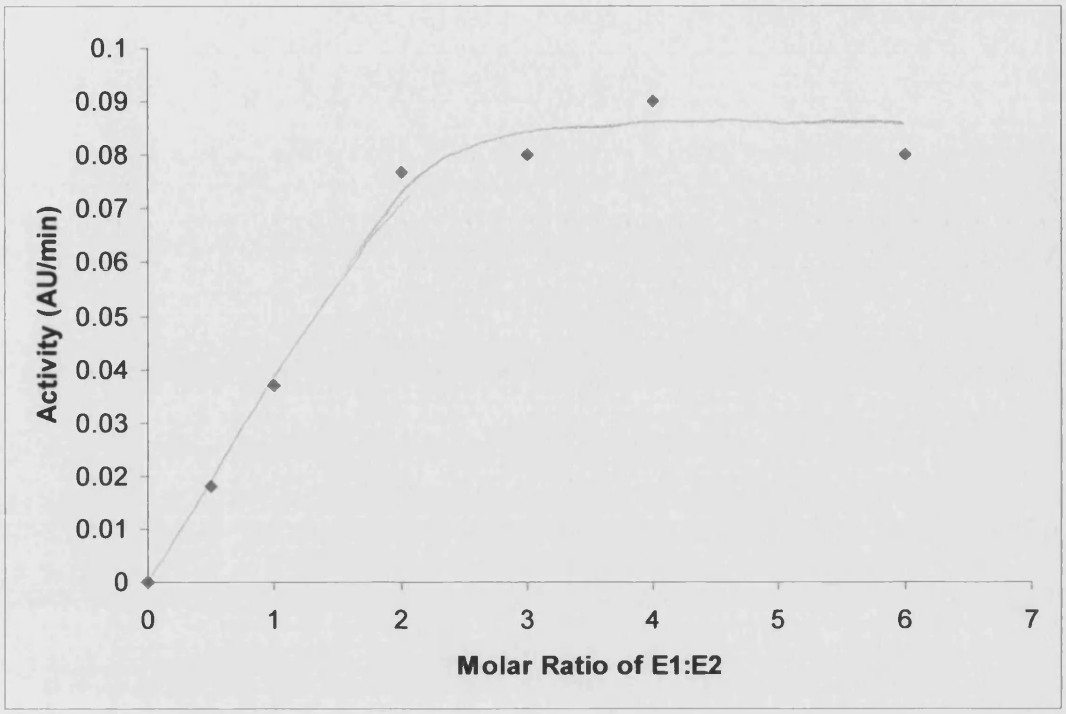


Fig.6.4. Graph showing complex activity as a function of varying molar ratios of E1. The ratio of E2 to E3 is fixed at 1:1.

Varying the molar stoichiometry of E3

A series of assays was performed, where the molar stoichiometry of E3 varied and the ratio of E1:E2 remained fixed. The stoichiometries of E1:E2:E3 were 3:1:0.05, 3:1:0.1, 3:1:0.25, 3:1:0.5 and 3:1:1. Complex was assayed for activity against the substrate 3-methyl-2-oxopentanoate. Complex activity was seen to decrease as the molar ratio of E3 to E1 and E2 mixed to assemble the complex increased from 0.05 to 1 (Fig. 6.5).

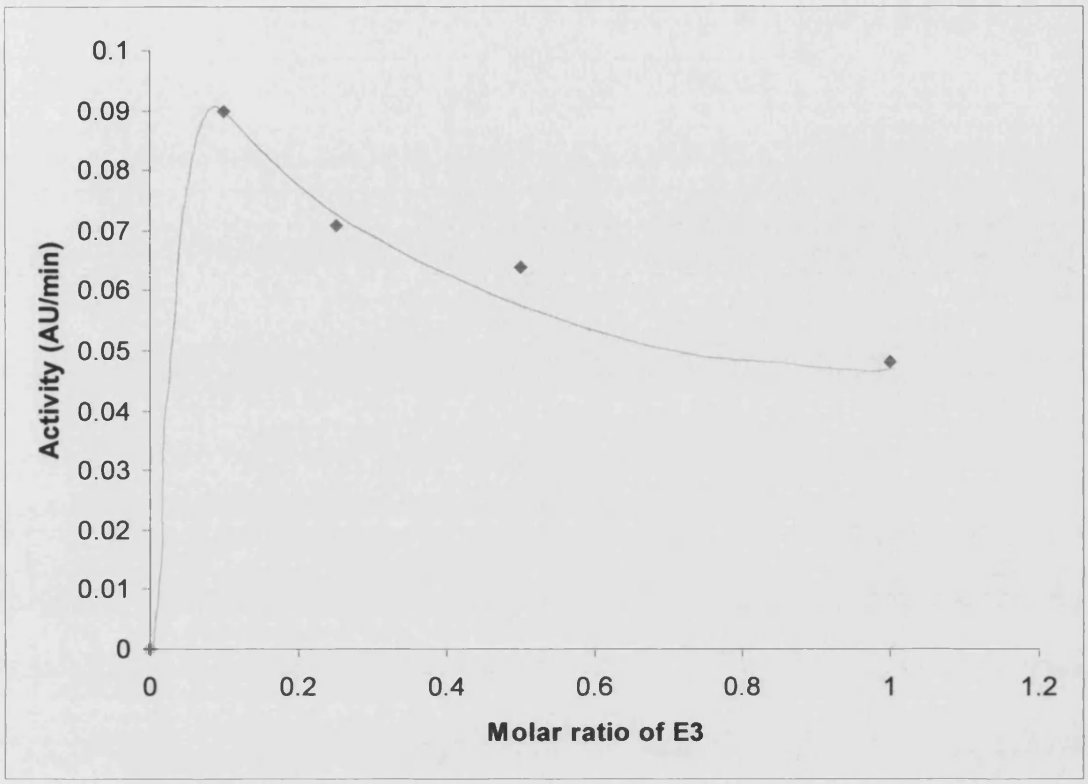


Fig. 6.5. Graph showing complex activity as a function of varying molar ratio of E3. The ratio of E1 to E2 is fixed at 3:1.

6.3.3.2 Pre-incubation experiments

Incubating enzymes separately

A series of experiments was performed, where complex was part-assembled by mixing two of the three component enzymes, in every possible combination, and the sub-complex incubated for 1 h at 55 °C. The third enzyme was then added, and the complete complex pre-incubated (in assay buffer containing TPP) for 10 min at 55 °C and assayed as normal. No difference in activity was seen with any of the combinations tested.

Varying incubation temperature and length

Assembled complex was incubated for either 1 h at 55 °C, 1 h at room temperature or overnight at 4 °C, and then assayed as normal (including the

standard 10 min, 55 °C pre-incubation in TPP-containing assay buffer immediately prior to assay).

Highest activity was observed when assembled complex was assayed immediately after the component enzymes were mixed. Incubating at room temperature for 1 h did not increase the activity, but incubating for 1 h at 55 °C and overnight at 4 °C both resulted in a 20 % decrease in rate.

6.3.3.3 Complex activity

The complex was active with the 3 branched chain 2-oxoacids 4-methyl-2-oxopentanoate, 3-methyl-2-oxopentanoate and 3-methyl-2-oxobutanoate. Activity was also observed with pyruvate as substrate, although this was 2.5-5 times lower than with branched-chain 2-oxoacids. No activity was observed with 2-oxoglutarate. An identical situation is observed in the BCOADHC of bovine kidney, where activity with pyruvate is 2.5-5 times lower than with the branched-chain 2-oxoacids (Pettit et al., 1978). The best specific activity value obtained was for the *Tp. acidophilum* BCOADHC with 3-methyl-2-oxopentanoate and was 4 U/mg E2. The specific activity ratios of the *Tp. acidophilum* and the bovine kidney complexes are compared in Table. 5.1, where values for the substrate 3-methyl-2-oxopentanoate were taken as 1 and substrate concentrations were saturating. The rate of production of NADH was directly proportional to the amount of enzyme in the assay, and there was a hyperbolic dependence of rate on the concentration of the one branched-chain 2-oxoacid substrate tested, 3-methyl-2-oxopentanoate. The K_m value was 250 µM (see figs. 6.6 and 6.7). This differs from the K_m value obtained with the same substrate for isolated E1 enzyme (Chapter 4), although the assay for E1 differs from the complex assay in that it utilises an unnatural electron acceptor in the removal of the acyl group from E1. Additionally, it might be that the active site of E1 is altered in some way when it is bound with E2 and E3.

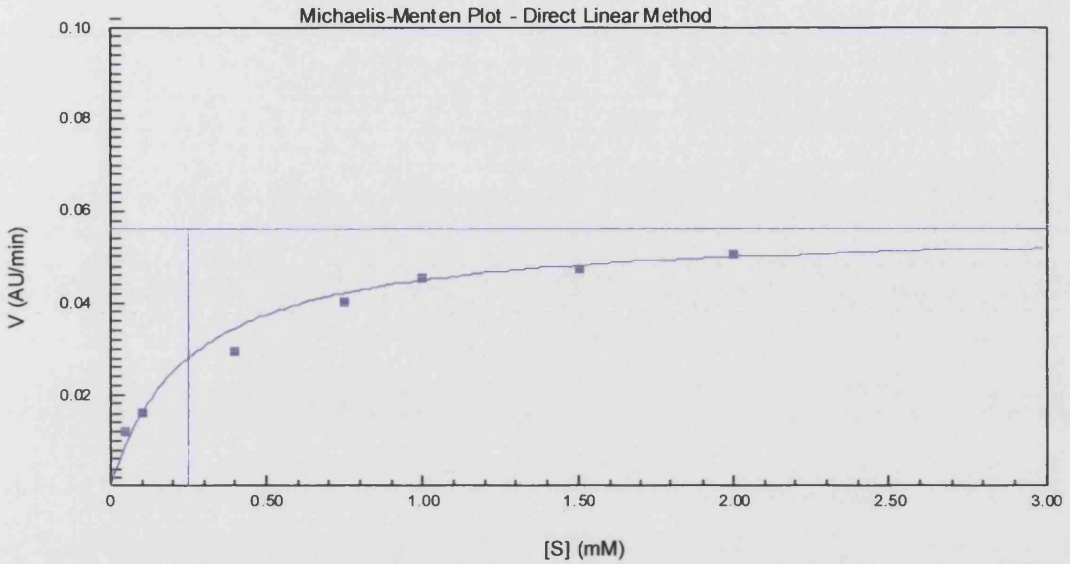


Fig. 6.6. The Michaelis-Menten plot of rate (Au/min) vs. [S] for 3-methyl-2-oxopentanoate.

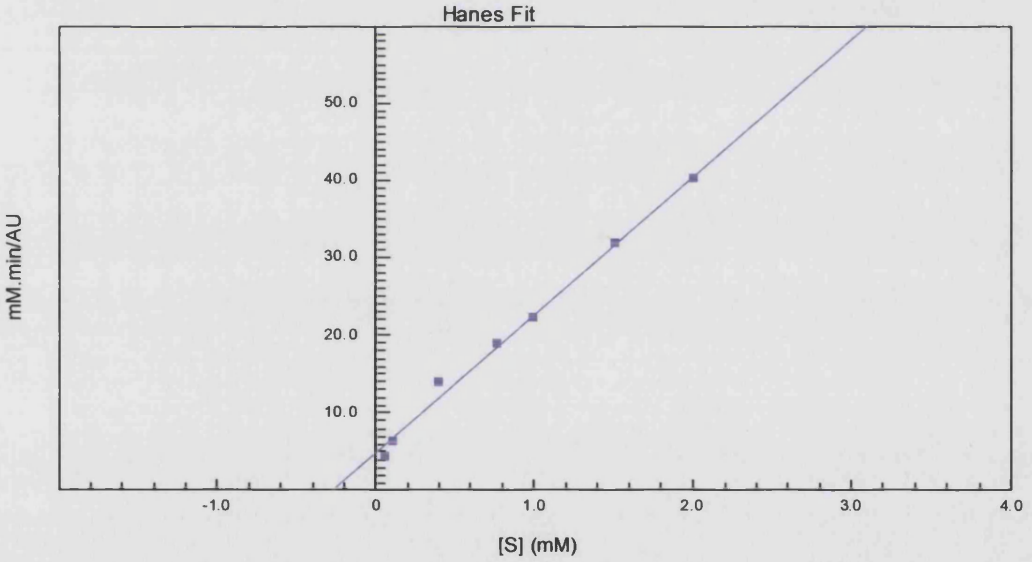


Fig. 6.7. Data obtained from E1 activity with 3-methyl-2-oxopentanoate are plotted using the Hanes Fit.

Table 6.1. Specific activity ratios of recombinant *Tp. acidophilum* BCOADHC, E1 and bovine kidney BCOADHC.

2-oxoacid substrate	Specific activity ratio:		
	Tp. acidophilum:		Bovine kidney
	BCOADHC	E1 (Ch. 4)	BCOADHC (Pettit et al., 1978)
3-methyl-2-oxopentanoate	1.0	1.0	1.0
4-methyl-2-oxopentanoate	0.5	0.3	1.5
3-methyl-2-oxobutyrate	0.9	0.6	2.0
pyruvate	0.2	0.2	0.4
2-oxoglutarate	0	0	0

6.3.3.4 OADHC purification by gel filtration

In preparation for analytical gel filtration complex was assembled in a 3:1:0.1 molar stoichiometry, and incubated for 10 min at 55 °C with 0.2 mM TPP. Analytical gel filtration revealed two major peaks, one representing material eluting at the column exclusion volume (> 1.3 MDa), and a second that corresponded with a size of around 160 kDa. A third, minor peak, corresponded with material of around 80 kDa in size (Fig. 6.8). Protein from the peaks was assayed for E1, E3 and complex activity, and subjected to SDS-PAGE (Fig. 6.9).

Protein in exclusion volume fractions was active in the OADHC assay. SDS-PAGE showed three bands corresponding to E1α (41 004 Da), E1β (37 445 Da) and E2/E3 (46 276 and 49 867 Da respectively; E2 and E3 run together on the gel). Protein in the fraction of 160 kDa was active in the E1 and E3 assays. Protein in the fraction of 80 kDa was active in the E3 assay. In repeat experiments, two peaks that correspond to exclusion volume and 160 kDa sizes are seen, but the height of the peaks varies with each run. Usually the third peak, containing E3 protein, is not seen. Also, in repeat experiments, the 160

kDa fractions contain only E1 α and β bands when analysed by SDS-PAGE and no other protein bands are present, unlike this example. Importantly in this context, OADHC activity is always seen only in the exclusion volume.

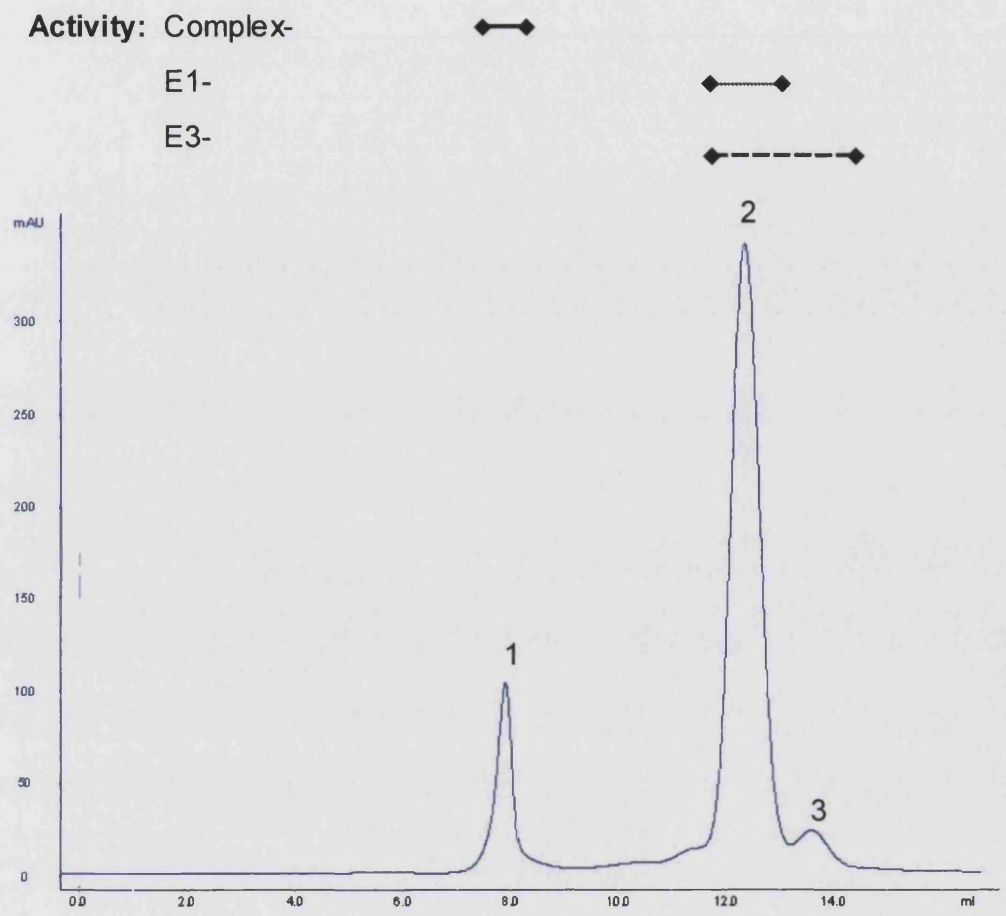


Fig. 6.8 (top). Gel filtration trace of OADHC. Protein in peak 1 is eluting at the column exclusion volume (> 1.3 MDa). Protein in peak 2 corresponds to a M_r of ~ 160 kDa and protein in peak 3 corresponds to a size of ~ 90 kDa. OADHC, E1 and E3 activities are indicated by bars above peaks.

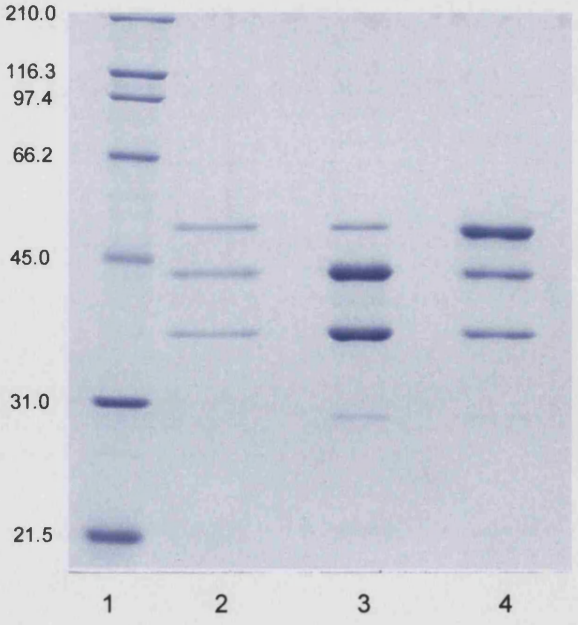


Fig 6.9 (left). SDS-PAGE of the three gel filtration peaks. Lane 1, standard protein markers (M_r values given in kDa); lane 2, protein from peak 1; lane 3, protein from peak 2; lane 4, protein from peak 3.

6.4 Discussion

This chapter of work demonstrates the *in vitro* assembly of an active BCOADHC from the recombinant E1, E2 and E3 protein products of *Tp. acidophilum* genes Ta 1438 (α), Ta 1437 (β), Ta 1436 and Ta 1435 respectively. The complex is active with the three branched-chain 2-oxoacids, and also with pyruvate, in ratios approximately the same as those observed with the isolated E1 enzyme. No activity was detected with 2-oxoglutarate. This has also been observed in the BCOADHC of bovine kidney, where activity with pyruvate was detected, but the specific activity was similarly 2-5 times lower than with branched-chain 2-oxoacids (Pettit et al., 1978). No oxidation of 2-oxoglutarate was observed in that complex. The specific activity of around 4 U/mg E2 for the *Tp. acidophilum* BCOADHC is comparable to the specific activity of the PDHC of *B. stearothermophilus* (8-10 U/mg E2) (Chauhan et al., 2000).

The E3 was cloned and expressed by Hans Christian Aass, and is reported in his thesis (Aass, 2003). In this chapter, gel filtration revealed that the E3 protein exists predominantly as an α_2 dimer, although a smaller amount of greater molecular weight species was revealed in the sample and was also active in the E3 assay. The E3 enzymes of other species (Mattevi, 1991; 1992 and references therein) similarly adopt an α_2 homodimeric conformation. The kinetic parameters of the *Tp. acidophilum* E3 are largely comparable with those of the recombinant E3 of *Haloferax volcanii* (Jolley et al., 1996).

The OADHC was assembled *in vitro* by mixing the individual recombinant E1, E2 and E3 subunits. Active OADHCs have previously been successfully assembled *in vitro* by other researchers (Lessard et al., 1998). Maximum activity of the *Tp. acidophilum* multienzyme complex was observed when the complex was assembled by mixing the individual subunits in a molar stoichiometry of 3:1:0.1 (E1:E2:E3). It should be stressed that the term 'stoichiometry' in this study refers to the manner in which the enzyme subunits were mixed *in vitro*, rather than referring to their actual association with each

other. However, given that in most complexes studied, the E1 step is rate-limiting, it makes sense that a stoichiometry of 3:1:0.1 gives optimal whole complex activity, and this value probably does reflect the molar stoichiometry of subunit assembly.

Activity increased in a hyperbolic manner as the complex was assembled with an increasing molar stoichiometry of E1 (from 1:1:1 to 6:1:1), with maximum activity occurring when subunits were combined at a molar stoichiometry between 2 and 3:1:1. Activity decreased, however, as the complex was assembled with an increasing molar stoichiometry of E3 (3:1:0.1 to 3:1:1). This suggests that E3 competes with E1 for binding to E2. In chapter 5 of this study, it was determined that the *Tp. acidophilum* E2 is 24-mer, and so probably has octahedral symmetry. The observations of competition between E1 and E3 in the current study is in agreement with the previous observations of other researchers, that the peripheral subunit binding domain of the octahedral BCOADHC E2 binds E1 or E3 mutually exclusively, and thus spatial competition between E1 and E3 for binding to the PSBD is witnessed (Lessard et al., 1998; Perham, 2000).

The native PDH complex of *B. stearothermophilus* has a substantial molar excess of E1 chains and a lower molar ratio of E3 to E2 (Henderson and Perham, 1980). In that example, the stoichiometry differed in three different preparations of the PDH complex. Reed et al. (1975) demonstrated that the native *E.coli* PDHC has an optimum stoichiometry for catalytic activity of 12 E1 and 6 E3 chains bound to a 24-mer E2 core. Only 18 of the 24 potential binding sites on E2 were filled. Bates et al. (1977) assembled the same *E. coli* PDHC *in vitro*, and achieved a subunit stoichiometry of 2:1:1. It is interesting to note that, in this case, an E2 subunit was able to bind both an E1 and E3 subunit, although whether this stoichiometry occurs in nature is not known. The PDHC of ox heart was found to have a stoichiometry of 1:1:0.16 (E1:E2:E3) (cited in Henderson and Perham, 1980). Perham (2000) explains that there is no fixed pattern to the

stoichiometry of any complex type, and there is great potential for structural isomerism. Additionally, a complex does not have to be fully assembled to be catalytically active.

Once a near-optimum molar stoichiometry for complex assembly had been established, the order in which subunits were combined together made no difference to activity. This feature was also observed for the *E. coli* PDHC (Reed et al., 1975).

Gel filtration showed that the complex elutes at the column exclusion volume, indicative of a size greater than 1.3 MDa. No complex activity was detected in peak fractions smaller than this. SDS-PAGE revealed three protein bands in this peak, which corresponded with E1α, E1β, E2 and E3. (the latter two bands run in an identical position because of their similar size, but both proteins must be present in order for the complex to be active in this peak). This indicates that the active complex has a M_r value greater than the association of a single E1, E2 and E3 subunit, which would be 300 kDa. The size of the E2 core was determined in Chapter 5 as ~ 1.1 MDa, indicative of an assembly of 24 E2 polypeptides. This is in agreement with the mammalian E2 core, which comprises 24 E2 chains (Pettit et al., 1978). Assuming that a full complement of E1 or E3 enzymes bind to the core, then the *Tp. acidophilum* OADHC would also be expected to have a size of 4-5 MDa.

In conclusion, the single *oadhc* operon in the archaeon *Tp. acidophilum* does indeed encode a functional OADHC, and the complex is a branched chain 2-oxoacid dehydrogenase. The complex spontaneously assembles *in vitro* from its individual enzyme components. It has a 24-mer E2 core, and so probably assembles to a size of ~ 5 MDa.

CHAPTER 7: ATTEMPTS TO INDUCE THE EXPRESSION OF OADHC ACTIVITY IN *THERMOPLASMA ACIDOPHILUM*

7.1 Introduction

Chapters 3-5 describe experiments that, together, generated a fully-active recombinant branched-chain 2-oxoacid dehydrogenase multienzyme complex from the putative *oadhc* operon in *Tp. acidophilum*. Whether or not the gene cluster is ever expressed *in vivo* remains to be established.

To address this question, *Tp. acidophilum* was grown in media containing either glucose, or the branched-chain amino acids L-valine and L-isoleucine, in an attempt to induce expression of the *oadhc* operon. These branched-chain amino acids were selected because of their ability to induce the *oadhc* operon of *P. putida* (Madhusudhan et al., 1999).

7.2 Methods

7.2.1 Growth of *Thermoplasma acidophilum*

Cells were cultured in DSM medium 158 with either glucose (1 % w/v), or the branched-chain amino acids valine and isoleucine (0.3 % and 0.1 %, respectively), as a carbon source.

Preparation of stock solutions

For one litre, four solutions were prepared and combined:

- *mineral salt medium (pH 1.6-1.7)*
- *trace elements stock solution*
- *Difco yeast extract (10 % w/v) stock solution*
- *glucose (50 % w/v) stock solution* (**or** *valine and isoleucine – these were not made as stock solutions, but were added as described below)*

- *Mineral salt medium.* 960 ml (980 ml for growth with amino acids) contained 1.32 g $(NH_4)_2SO_4$, 0.372 g KH_2PO_4, 0.247 g $MgSO_4.7H_2O$ and 0.074 g $CaCl_2.2H_2O$. The pH was adjusted to pH 1.6 - 1.7 at room temperature with concentrated H_2SO_4, and the solution was then autoclaved.

- *Trace element solution.* A stock solution of 1 litre contained 1.93 g $FeCl_3.6H_2O$, 0.18 g $MnCl_2.4H_2O$, 0.45 g $Na_2B_4O_7.10H_2O$, 22 mg $ZnSO_4.7H_2O$, 5 mg $CuCl_2.2H_2O$, 3 mg $Na_2MoO_4.2H_2O$, 3.8 mg $VOSO_4.5H_2O$ and 2 mg $CoSO_4.7H_2O$. The pH of this solution was not adjusted. It was autoclaved.

- *Difco yeast extract.* A 10 % (w/v) stock solution was made and filter sterilised, rather than autoclaving.

- *D-glucose.* A 50 % (w/v) stock solution was made (heating the mixture by placing in the 59 °C incubator assisted in dissolving) and filter sterilised.

Preparation of medium

-containing glucose. To the 960 ml of mineral salt media, 20 ml glucose, 10 ml yeast extract and 10 ml trace elements from the stock solutions were added immediately prior to inoculation. The medium was heated to 59 °C in the incubator before inoculation.

-containing branched-chain amino acids. To replace glucose with valine and isoleucine, mineral salt medium was made as before, but to a volume of 980 ml. Around 40 ml was decanted from this, and used to dissolve 3 g L-valine and 1 g L-isoleucine separately. These solutions were filter sterilised and returned to the medium, and other ingredients added as before, to give 1 litre.

Often, once all components of the growth medium were mixed together, a brownish precipitate could be seen. This had no effect on the growth of the cells, and the medium was never discarded if the precipitate formed, as during shaking, the precipitate appeared to re-dissolve.

Inoculation

Cultures were initiated from a dry pellet supplied by DSMZ (Braunschweig, Germany) following the accompanying instructions. The small volume starter-cultures were grown in media containing glucose, and then transferred to media containing branched-chain amino acids once culture volumes increased to 500 ml.

DSM 1728 dry cells were resuspended in 1 ml warm medium, over a 30 min period. 0.5 ml was then transferred to 10 ml warm medium in a capped glass tube, and incubated at 59 °C without shaking, on its side (0.5 ml was retained at 4 °C for future use). When the cell density had reached A_{600} ~ 0.12, cells were transferred to fresh, warm medium. Routinely, a 10 % inoculum was used, and in this way the culture volume was increased to 500 ml in 2 litre, uncapped flasks. The desired density for subculture (A_{600} = 0.1-0.2) was usually reached on the third day. Subculture was carried out on the fourth day, and cells were not left for any longer than this. Once the culture volume had reached 500 ml, cultures were incubated with shaking at 100 rpm. Cells were harvested at A_{600} = 0.2, by centrifugation at 13000 × *g*, 4 °C for 15 min and stored at -20 °C.

7.2.2 Preparation of soluble extracts

Frozen cells were resuspended in 50 mM Tris buffer, pH 8.0 (800 µl buffer/200 mg cells) with 1 mM PMSF and 1µl benzonase nuclease, and incubated on ice for 20 min. The soluble extract was then separated from cell debris by centrifugation at 13000 × *g* and retained on ice for immediate use in enzyme assays. Protein concentration was determined by the method of Bradford (1976).

7.2.3 OADHC and E3 assays

Cell extracts of *Tp. acidophilum* cells grown in the presence of either glucose or the branched-chain amino acids were assayed for OADHC and E3 activity. The assays were carried out as described in Section 2.5. Up to 200 µl of cell extract was used in the assays.

7.2.4 Storage of cells and re-initiation of a growing culture

A growing cell suspension of A_{600} 0.1-0.2 was placed at 4 °C with a sealed lid, and stored for up to 4 months. The culture was re-initiated by slow reheating to 59 °C as follows; the culture was placed at room temperature for 1-3 days, and then transferred to a cool incubator that was then turned on and set at 59 °C. Cells heated to 59 °C slowly as the incubator temperature rose. After 24 h, the culture was checked for motility under the microscope. It was then used to inoculate fresh, glucose-containing medium (the inoculum was 10 % of the total medium volume).

7.3 Results

7.3.1 Growth

When growth was first initiated from dry pellets, cultures of a small volume were grown without aeration. However, once culture volumes increased to 500 ml, cultures were aerated by shaking at 100 rpm. This appeared to enhance growth slightly, with cells appearing healthy and highly active under the microscope, and reaching an A_{600} of 0.1- 0.2 on the third day after inoculation. Subculturing was then carried out on the fourth day with an inoculum that was 10 % of the final new culture volume.

Around 300 mg wet cell pellet was obtained from a 1 litre culture harvested on the fourth day of growth, whether grown in glucose or branched-chain amino acid-supplemented medium.

7.3.2 Enzyme assays

Preparation of cell extract

When a cell extract was prepared as detailed in Section 7.2.2, a preparation containing between 4 and 19 mg protein /ml could be obtained.

Activity

Up to 200 µl of cell extract was used in the OADHC and E3 assays. The substrate in the OADHC assay was 3-methyl-2-oxopentanoate. Neither OADHC nor E3 activity was detected in cell extracts of cells grown with either carbon source. The system employed for assay is able to detect activity at levels of 3 x 10^{-4} U/mg.

7.4 Discussion

Activity

The previous chapters of the current study revealed that a putative *oadhc* operon in *Tp. acidophilum* encodes a BCOADHC. The aim of this chapter of work was to explore whether the complex is present and functional *in vivo*. The approach taken was to attempt to induce *oadhc* expression by growing the cells in medium that contained either L-branched-chain amino acids or glucose as the sole carbon source, and then examining cell extracts from both cultures for BCOADHC activity.

No activity was detected in any of the *Tp. acidophilum* cell extract preparations. Other groups have detected BCOADHC activity in cell extracts from *P. putida* and bovine kidney, at levels of 0.04 – 0.1 U/mg protein (Marshall and Sokatch, 1972; Pettit et al., 1978; Sokatch et al., 1981). The system used in the current study can detect an appearance of NADH at 3×10^{-4} U/mg protein, and so if any BCOADHC activity was present in the *Tp. acidophilum* cell extracts, it would have been at levels < 1 % of the other BCOADHCs.

Because so many factors are involved in the regulation of gene expression (Alberts et al., 1994), only a few can be considered here. As stated in Chapter 1, the putative *oadhc* operon of *Tp. acidophilum* does have a TA-rich potential promoter region upstream of the first open reading frame that could serve as a transcriptional promoter, and so it has the potential to be transcribed. It is possible that the operon is not induced under the growth conditions to which the cells were exposed, or that other transcription regulatory factors, such as those associated with catabolic repression (CR), are involved. Additionally, it could be that essential transcription factors are defective or missing, or that regulation is occurring at a translational level. It might be that the complex is expressed and is present in cells, but is inhibited, is not functional, or dissociates during the preparation of cell extracts prior to assay.

Regulation of gene expression

For cells grown with glucose, it is possible that catabolic repression occurs. CR is a mechanism where the expression of genes involved in the metabolism of a growth substrate is inhibited by a more readily available, metabolisable carbon source. Transcription of the *bcoadhc* (encoding a BCOADHC) of *P. putida* is repressed by glucose, succinate and ammonium ions, as shown by addition of those catabolites to a valine-isoleucine growth medium (Sykes et al., 1987) and the *bcoadhc* of *Enterococcus faecalis* is similarly repressed by glucose, as well as fructose and lactose. In the latter example, the upstream region of the first open reading frame contains a CR sequence TGTATGCGCTTACA, which overlaps the -35 region of the promoter (Ward et al., 2000). This was searched for, but is not present in the *Tp. acidophilum* operon. Of course, the archaeal CR sequence could be very different from this bacterial example, and so catabolic repression cannot be ruled out as the reason for the absence of activity in cells grown in medium containing glucose. One would imagine, however, that if this were the case, then activity would be observed in cells grown with branched-chain amino acids, but this was not the case.

Perhaps the operon is not induced with the L-branched-chain amino acids. The BCOADHC of *P. putida* has been studied extensively by the group of John Sokatch at the University of Oklahoma, USA. The group states that it is the L-branched-chain amino acids that are the inducers of the operon (Madhusudhan et al., 1999 and references therein). The *bcoadhc* promoter region of DNA was cloned into a plasmid vector, upstream of a promoterless *lacZ* gene. β-Galactosidase activity was detected when cells were grown with L-valine (0.3 %) and L-isoleucine (0.1 %). In earlier work, the group had hypothesised that the branched-chain 2-oxoacids, and not the branched-chain amino acids, were the inducers (Marshall and Sokatch, 1972), but that hypothesis was revised, and currently the group claim that induction of the operon in cells grown in medium with the branched-chain 2-oxoacids must be due to conversion of these

compounds to their respective L-branched-chain amino acids (Madhusudhan et al., 1999).

Yet, in contrast to the theories of Sokatch and co workers, Ward et al. (1999; 2000) state that it is, in fact, the 2-oxoacids, and not the branched-chain amino acids that are responsible for *bcoadhc* induction in *E. faecalis*. The researchers carried out studies on induction and repression of the *E. faecalis bcoadhc* promoter. The promoter was cloned as a transcriptional fusion to the β-glucuronidase (GUS) gene in *E. coli* plasmid pDW100. The effects of the branched-chain amino acids, the branched-chain 2-oxoacids, glucose and fructose on transcription were then assessed by measuring GUS activity. The studies provide strong evidence that the BCOADHC promoter of *E. faecalis* is induced by the branched-chain 2-oxoacids rather than the branched-chain amino acids. Attempts were not made in the current study to induce the *Tp. acidophilum oadhc* by growth with the branched-chain 2-oxoacids. This is an important consideration for future experiments.

The transcriptional activator BkdR is required for transcription of the *P. putida bcoadhc* (termed the *bkd* operon by Madhusudhan et al., 1999). This activator is a homologue of the *E. coli* Lrp (leucine-responsive protein), with 37 % sequence identity and 58 % similarity to Lrp, but BkdR differs from that global example in that it strictly serves only as activator of the *oadhc* operon. Deletions in even a few bases in the 5' region of the BkdR binding site results in a great reduction in transcription (Madhusudhan et al., 1999 and references therein). A possible transcription factor was searched for in *Tp. acidophilum* genome, using the *P. putida* BkdR sequence (GI 790513) for comparison, but no significant matches were found. Unless a different type of transcription factor is utilized other than BkdR, then perhaps the operon does not rely on the action of transcription factors for expression. Alternatively, the absence of a functional transcription factor may be the reason for the lack of activity.

It is possible that regulation of *oadhc* expression is occurring at the level of translation. Bacterial and eukaryotic translational regulation can involve many factors, including control of translational initiation and degradation of mRNA (Alberts et al., 1994). Little is known about translational regulation or the stability of mRNA species in archaea. Jager et al. (2002) have, however, investigated whether mRNA processing contributes to regulation of gene expression in *Haloferax mediteranei*, by analysis of the decay of mRNA species under different salt concentrations in the growth media. They concluded that it is indeed the case that mRNA degradation is a method of gene regulation in archaea, but they state that much more research into this area is required. It is possible that the *Tp. acidophilum* is transcribed, but then regulated at the translational level. Indeed, an mRNA species for the OADHC of a halophilic archaeon has been detected by Northern analysis, and subsequently by RT-PCR, and yet no OADHC enzymic activity has been detected in that organism (Jolley et al. 2000, Al-Mailam, 2006).

The complex is inhibited, is not functional, or has dissociated

It was considered that the *Tp. acidophilum* BCOADHC might be regulated by phosphorylation. The genome of *Tp. acidophilum* was searched for phosphatases and protein kinases, using blastp for bovine kidney (GI 27806819) and human (GI 45439339) pyruvate dehydrogenase phosphatase sequences (no branched-chain dehydrogenase phosphatase sequences were available), along with mouse (GI 2809220) and human (GI 21431746) branched-chain 2-oxoacid protein kinase sequences. Default parameters were used in the search. The genome does not appear to contain either of these regulatory enzymes, and so the complex is probably not inactivated by a system of this type.

It is unlikely that the absence of activity is due to dissociation of the multienzyme complex during the preparation stages of soluble extract. The method of cell lysis for *Tp acidophilum* is very gentle. The cells lyse readily when resuspended in 50 mM Tris (pH 8.0) buffer. Moreover, the complex is stable under the

conditions used in the recombinant experiments of Chapter 6. Additionally, no E3 activity was detected in any of the cell extracts, whereas activity would have been expected should it be that the complex had been present, but was inhibited or had fallen apart.

Interestingly, E3 activity has previously been detected in cell extracts of *Tp. acidophilum* (Smith et al., 1987) and also of the halobacteria (Danson et al., 1984); in the latter case, complex activity was also tested for and was absent. Why activity of that one complex enzyme can be detected remains to be explained. Perhaps, on those occasions, activity was a result of background expression. When an operon is repressed, the repressor protein molecule usually attaches to the DNA at the region of the promoter. However, the state of repression of a gene will exist in equilibrium with a state of non-repression (see Fig. 7.1). Even though the repressed state is favoured, this equilibrium does allow a small degree of transcription to occur. As E3 is a more active enzyme than the other complex enzymes (with specific activity values of 17 U/mg for E3, compared to 0.4 U/mg for E1, in recombinantly-produced enzymes of *Tp. acidophilum*), this might explain why only activity of this component is detected.

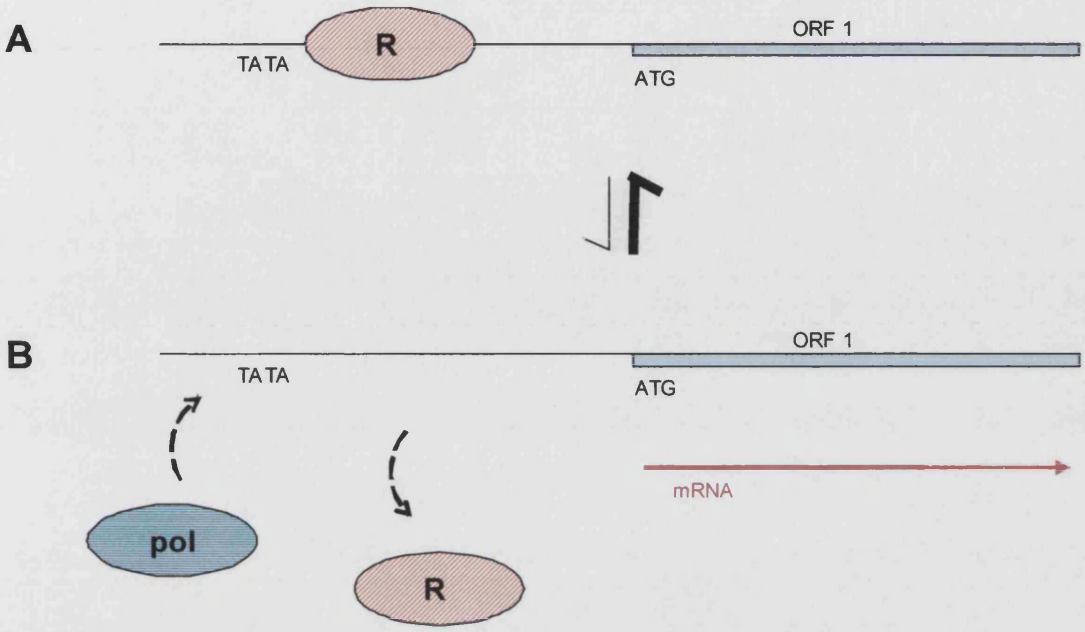


Fig. 7.1. An equilibrium exists between a repressed state (A) and a non-repressed state (B). Although state A is favoured, a small amount of transcription is able to occur when the repressor protein dissociates from the promoter.

It is unlikely that only the E3 enzyme of the complex is functional. Recombinant studies have already shown that the E1 genes encode functional branched-chain decarboxylase (Chapter 4). However, the *Tp. acidophilum* genome was searched for evidence of branched-chain amino acid transaminase enzymes (using blastp sequence comparisons with *Bacillus licheniformis* branched-chain aminotransferase II (GI 52004526), *Pseudomonas putida* branched-chain aminotransferase (GI 26990223) and human branched-chain aminotransferase I (GI 38176287)), which would generate the branched-chain 2 oxoacid substrates from the branched-chain amino acids. None were identified. Additionally, no branched chain amino acid transporters were found in the genome (using blastp sequence comparison with *P. putida* branched-chain amino acid transporter GI 8273925). No genomes in the NCBI database appeared to have 2-oxoacid

transporters. This raises questions as to how the 2-oxoacid substrates are made available to the complex.

To conclude, it is possible that either the *Tp. acidophilum oadhc* operon is not induced under the growth conditions detailed in this chapter, or that regulation is occurring at other post-transcriptional levels. If the complex had been present, whether in very low amounts or in a dysfunctional form, then detectable E3 activity would have been expected, and this was not the case.

CHAPTER 8: DISCUSSION AND CONSIDERATIONS FOR FUTURE WORK

8.1 The discovery of a functional archaeal OADHC

This work has shown that the putative *oadhc* operon in *Thermoplasma acidophilum* does indeed encode a functional 2-oxoacid dehydrogenase multienzyme complex, and that the complex is a branched-chain 2-oxoacid dehydrogenase. This is the first report of a functional OADHC in the Domain Archaea.

The recombinant complex probably resembles the putative native complex; it has an octahedral 24-mer E2 core and a 165 kDa E1 of $\alpha_2\beta_2$ conformation, like its eukaryotic and bacterial counterparts. Furthermore, kinetic analysis reveals that the branched-chain 2-oxoacids are probably the true substrates for the complex and kinetic values are comparable to eukaryotic and bacterial BCOAHDCs and PDHCs (Pettit et al., 1978; Lessard and Perham, 1994). The complex also has activity with pyruvate, but this is at lower levels than the branched-chain 2-oxoacids and so it is not clear whether or not the BCOADHC might have a physiological role in the degradation of pyruvate *in vivo*. The results suggest that the archaeal BCOADHC should be active *in vivo*, but no activity was detectable in cell extracts of cells grown with either glucose or the branched-chain amino acid precursors to the branched-chain 2-oxoacid substrates. The growth studies revealed that the complex is either not present, or is not active *in vivo*. Thus it cannot yet be said whether archaea actually utilise OADHCs in their central metabolic pathways.

OADHCs are active in aerobic bacteria, and in the mitochondria of eukaryotic cells (Danson et al., 2004 and references therein). It was formerly believed that

these large complexes evolved first in aerobic members of Bacteria *after* the Bacteria-Archaea evolutionary divide, and that eukaryotes acquired them from the bacterial ancestor to the mitochondrion, as a result of endosymbiosis (Schnarrenberger and Martin, 2002). The identification of putative *oadhc* operons in archaeal species, and the discovery that an archaeal OADHC is active now calls for consideration of how these genes arose in Archaea. The options are that either extensive lateral gene transfer has occurred between bacteria and archaea, or the *oadhc* genes were in fact present in an ancestor common to those two Domains.

It is unlikely that the presence of putative OADHC genes in certain archaeal species is a result of random lateral gene transfer events between those species and bacterial ones, as is claimed by Wanner and Soppa (2002). The putative genes are found in all aerobic archaeal species examined to date, from both Crenarchaeota and Euryarchaeota, and the genes have evolved such that their protein products would be adapted to the environment in which each species lives. For example, the putative halophilic OADHC genes have the high GC content characteristic with those genomes from high salt environments. It seems reasonable to suggest that an OADHC may have arisen in an ancestor to Bacteria and Archaea and has been retained in aerobic species of these Domains. In Bacteria, the complexes may have diverged to give the different family members with different substrate specificities, whereas the situation in Archaea is currently unknown, as the substrate specificities have not yet been elucidated.

8.2 A putative role for an archaeal OADHC

The putative operon has been identified in all aerobic archaeal species whose genomes have been sequenced, although the actual substrate specificities for these organisms have not yet been determined. Only the complex type present in *Tp. acidophilum*, a BCOADHC, is now known. This is the only putative *oadhc* operon in the genome of this organism. Although *Hfx. volcanii* has a complete

putative operon, the substrate specificity for that complex will need to be determined through recombinant expression of the genes, because no activity can be detected in those cells either (Dina Al Mailam, personal communication). The presence of such a huge multienzyme complex in an archaeon is somewhat of a mystery, whatever its family type and role may be.

The roles of both pyruvate and 2-oxoglutarate dehydrogenases are in ATP generation (with both enzyme complexes functioning within the central metabolic pathway). The branched-chain 2-oxoacid dehydrogenases are reported to have several roles, including ATP generation in *P. putida*, an involvement in cell-cell signalling in the spore-forming bacterium *Myxococcus xanthus* and a function in the production of the branched-chain fatty acid precursors for membrane synthesis in many bacteria (Sokatch et al., 1981; Oku and Kaneda, 1988; Downard and Toal, 1995). In *M. xanthus*, an essential signal in cell-cell communication, termed the E-signal, is required for expression of many developmental genes involved in the formation of fruiting bodies and spores. The role of the BCOADHC in this species is the generation of the branched-chain fatty-acids that constitute the E-signal (Downard and Toal, 1995). However, there is no analogous sporulation in archaea. The BCOADHC of *Streptomyces avermitilis*, a Gram positive soil bacterium that produces anthelmintic avermectins, has an essential role in the production of the branched-chain fatty acid precursors of the polyketide backbone of this type of antibiotic (Denoya et al., 1995). Again, in archaea there is no polyketide synthesis. Archaeal membrane lipids do not contain fatty acids, instead comprising isoprenoid lipids. It is perhaps a possibility that *Tp. acidophilum* utilises a BCOADHC for synthesis of the acyl-CoA precursors of isoprenoid lipids (Boucher et al., 2004).

As was explained in the introduction (Chapter 1), archaea have a family of 2-oxoacid ferredoxin oxidoreductases (FORs) that carry out the catabolic conversions of 2-oxoacids to the corresponding acyl-CoAs. The family is akin to

that of the OADHCs, in that members are capable of decarboxylating pyruvate, 2-oxoglutarate or the branched-chain 2-oxoacids, depending upon the type of FOR. However, some FORs have wider substrate specificity and can act upon either pyruvate or the branched-chain 2-oxoacids, for example, making it hard to say whether a particular organism would be expected to have the genes for all 3 FOR types in its genome, or whether an OADHC might replace a corresponding FOR and carry out the equivalent catabolic step. *Tp. acidophilum* has one complete *for* operon in its genome, but each putative gene is annotated differently, as either a pyruvate or 2-oxoacid FOR component, and so the substrate specificity for this enzyme is unclear. Furthermore, *Tp. acidophilum* has never been assayed for FOR activity. Whether or not the organism contains an FOR that can decarboxylate pyruvate, 2-oxoglutarate and/or branched-chain 2-oxoacid might give clues to the authenticity, or the possible role of its putative BCOADHC, be it a replacement for a branched-chain FOR, or whether it has an additional role, not related to ATP generation. FOR assay is an important consideration for future work.

8.3 Further considerations

The putative *oadhc* operon in *Tp. acidophilum* encodes a functional BCOADHC, and yet its activity cannot be detected *in vivo*. Future research needs to assess whether or not the complex is present but inactive, or if it is transcribed at all.

RT-PCR studies of the putative *oadhc* of *Hfx. volcanii* have revealed an mRNA species of the expected 5 kb operon size (Al Mailam, 2006), and this information would also be useful for the case of *Tp. acidophilum*. As no promoters for internal genes can be identified for *Tp. acidophilum oadhc*, only that which is located upstream of E1α, one would expect that a single mRNA transcript would be produced for the putative operon.

To examine *oadhc* expression further, immuno-techniques could be applied that use antibodies raised against *Tp. acidophilum* OADHC components. Also,

modern proteomics is proving a useful tool in examining gene expression, and has been applied in the investigation of the expression of genes encoding metabolic enzymes in *Sulfolobus solfataricus* (Snijders et al., 2006).

Further growth studies

Tp. acidophilum was grown with branched-chain amino acids, in an attempt to induce the complex, but this did not result in any detectable OADHC activity. There is the potential to carry out more extensive growth studies, including growth with the 2-oxoacid substrates, or combinations of sugar sources. In the current study, cells were only grown with branched-chain amino acids as a sole carbon source for as long as was required to obtain sufficient cell extract for assay, so their longevity in this type of medium was not determined. If the BCOADHC functions in a role not associated with central metabolism, then the cells may still need glucose as a carbon source. Additionally, if the transaminase enzymes or amino acid transporters really are absent (as appears to be the case, from genome analysis) and the 2-oxoacids are the inducers of the complex, then cells may need to be supplied directly with the branched-chain 2-oxoacids in order for the *oadhc* to be induced. Assay for transaminase activity (Cooper et al., 2002) might also be considered.

Extensive growth studies have, in fact, been carried out on *Hfx. volcanii*, where the E3 gene of the putative OADHC was inactivated by insertional mutagenesis. However, no change in phenotype, from that of the wild type, could be detected when mutants were grown on a range of sugars, amino acids and metabolites (Jolley et al., 1996).

It might also be a consideration to grow the cells anaerobically. The group of Jorg Soppa discovered that the E1 α and β products of a partial, putative *oadhc* gene cluster in *Hfx. volcanii* has an important function in anaerobic, but not aerobic, oxidation of some component of Casamino acids. They carried out growth studies on a nitrate respiration-deficient mutant which, by

complementation, had been shown as deficient solely for the *oadhc* genes. The mutant was able to grow only half as well as the wild type on Casamino acids, under anaerobic conditions where nitrate was an alternative electron acceptor (Wanner and Soppa, 2002). It is possible to grow *Tp. acidophilum* anaerobically if elemental sulphur is provided as an alternative electron acceptor and so perhaps the BCOADHC has a similar function in anaerobic respiration. Future work might involve mutation of the *Tp. acidophilum oadhc*, and anaerobic growth studies to elucidate whether this results in a change in phenotype, akin to experiments of that group.

The Soppa group has also studied the regulation of proteins involved in central metabolism in *Hfx. volcanii*, using a shotgun approach DNA microarray (Zaigler et al., 2003). They were able to assess which genes were up- or down-regulated when growing cells were transferred from Casamino acids to glucose as a sole carbon source. They discovered that a set of OADHC genes were repressed by glucose, as was an oligopeptide transporter protein. However, a second OADHC was induced, indicating that *Hfx. volcanii* might have more than one functional OADHC. Three putative operons / part operons have already been discovered in that partially-sequenced genome (Zaigler et al., 2003). As the entire genome of *Tp. acidophilum* has been sequenced, a whole genome array could be carried out.

Research into *Tp. acidophilum* lipoylation enzymes

Lipoylation of the lipoyl domain of E2 is essential for correct functioning of multienzyme complexes. As was described in the introduction, in *E. coli* there are two discrete pathways of achieving lipoylation, exogenous and endogenous, where lipoic acid is either taken up from the external environment, or synthesised intracellularly, respectively. In the exogenous pathway, lipoylation is carried out by lipoate protein ligase (lplA). The putative *lplA* gene of *Tp. acidophilum* (Ta 0154) has recently been cloned and expressed in *E. coli* (McManus et al., 2005). The protein product, a monomer of 29 kDa, was

catalytically inactive. Further examination of the gene sequence, compared to the *lplA* genes of *E. coli* (Gram-negative) and *Streptococcus pneumoniae* (a Gram-positive bacterium), revealed that the *Tp. acidophilum lplA* gene is around 270 bp shorter than its bacterial counterparts, and in fact, lacks the region which encodes an entire C-terminal domain in those enzymes. The absence of this C-terminal domain in the *Tp. acidophilum* enzyme might be the feature that renders it inactive. Interestingly, a gene (Ta 0513) that encodes a protein with structural homology to the C-terminal domains of the lplA enzymes of *E. coli* and *S. pneumoniae* has been located in the *Tp. acidophilum* genome, directly upstream of the N-terminal encoding region of the *lplA* gene (McManus et al., 2005). The genes overlap by one nucleotide, so that the A of the TGA stop codon of Ta 0513 is also that of the ATG the start codon of Ta 0514. It is intriguing that this C-terminal encoding gene is upstream of the 5' terminus of *lplA* gene, and that the feature is also witnessed in *Tp. volcanium* and *Halobacteria* (Mareike Posner, personal communication).

McManus et al. (2005) hypothesise that the Ta 0513 protein product might be required in association with Ta 0514 for a complete functional lplA enzyme. Kim et al. (2005) crystallised the Ta 0514 protein product, and incubated crystals first with ATP and subsequently lipoic acid. They observed a lipoyl-AMP intermediate bound in the active site of the Ta 0514 protein. Perhaps the two proteins combined can catalyse a two-step reaction, the first being the formation of the lipoyl-AMP intermediate, and the second, the transfer of the lipoyl group from it to the lysine residue target for lipoylation in E2.

An important goal for future work will be to establish whether *Tp. acidophilum* has a functional lplA, because deficiency of this enzyme will render the multienzyme complex inactive. Cloning and expression of the Ta 0513 gene has now commenced, as part of a subsequent Ph.D. project. It may be that the two genes require co-expressing in order to obtain correctly folded, soluble protein, as was the case for the E1α and β genes in this study.

The evidence put forth by this project and by previous research in this laboratory strengthens the suggestion that archaea possess a functional OADHC. Future work now needs to focus upon detecting multienzyme complexes *in vivo* and defining their metabolic function.

References

AASS, H.C. (2003) Multienzyme Complexes in the Archaea. *Ph.D. thesis report.*

AEVARSSON, A., SEGER, K., TURLEY, S., SOKATCH, J. R. and HOL, W. G. J. (1999) Crystal structure of 2-oxoisovalerate and dehydrogenase and the architecture of 2-oxo acid dehydrogenase multienzyme complexes. *Nature Structural Biology,* 6, 785-792.

ALBERTS, B., BRAY, D., LEWIS, J., RAFF, M., ROBERTS, K., WATSON, JD. (1994) *Molecular Biology of the Cell.* 3rd edition. Garland Publishing, New York.

Al-MAILAM, D (2006) The 2-oxoacid dehydrogenase complex of *Haloferax volcanii. Ph.D. thesis report.*

AMANO, N., TSUJI, K., EBIHARA, S. and SUZUKI, M. (2003) Identification of promoter sequences using archaeal genomic sequences - 2. Analysis of a meso-thermophile, *Thermoplasma volcanium. Proceedings of the Japan Academy Series B-Physical And Biological Sciences,* 79, 163-169.

ARJUNAN, P., NEMERIA, N., BRUNSKILL, A., CHANDRASEKHAR, K., SAX, M., YAN, Y., JORDAN, F., GUEST, J. R. and FUREY, W. (2002) Structure of the pyruvate dehydrogenase multienzyme complex E1 component from *Escherichia coli* at 1.85 angstrom resolution. *Biochemistry,* 41, 5213-5221.

BARNS, S.M., DELWICHE, C.F., PALMER, J.D., PACE, N.R. (1996) Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proceedings of the National Academy of Sciences of the United States of America.* 93, 8797-9.

BATES, D. L., DANSON, M. J., HALE, G., HOOPER, E. A. and PERHAM, R. N. (1977) Self-assembly and catalytic activity of pyruvate-dehydrogenase multienzyme complex of *Escherichia coli. Nature,* 268, 313-316.

BELL, S. D. and JACKSON, S. P. (1998) Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features. *Trends in Microbiology,* 6, 222-227.

BELL, S. D. and JACKSON, S. P. (2001) Mechanism and regulation of transcription in archaea. *Current Opinion in Microbiology,* 4, 208-213.

BERG, J.M., TYMOCZKO, J.L., STRYER, L. (2001) *Biochemistry.* Fifth Edition. W.H. Freeman, New York.

BOUCHER, Y., KAMEKURA, M. & DOOLITTLE, W. F. (2004) Origins and evolution of isoprenoid lipid biosynthesis in archaea. *Molecular Microbiology,* 52, 515-527.

BRADFORD, M.M. (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry,* 72, 248-254.

BROCHIER, C., GRIBALDO, S., ZIVANIVOC, Y., CONFALONEIRI, F., FORTERRE, P. (2005) Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to *Thermococcales*? *Genome Biology*, 6, R42.

CATE, R., ROCHE, T. and DAVIS, L. (1980) Rapid intersite transfer of acetyl groups and movement of pyruvate dehydrogenase component in the kidney pyruvate dehydrogenase complex. *Journal of Biological Chemistry,* 255, 7556-7562.

CHANG, C. F., CHOU, H. T., CHUANG, J. L., CHUANG, D. T. and HUANG, T. H. (2002) Solution structure and dynamics of the lipoic acid-bearing domain of human mitochondrial branched-chain alpha-keto acid dehydrogenase complex. *Journal of Biological Chemistry,* 277, 15865-15873.

CHAUHAN, H. J., DOMINGO, G. J., JUNG, H. I. and PERHAM, R. N. (2000) Sites of limited proteolysis in the pyruvate decarboxylase component of the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus* and their role in catalysis. *European Journal of Biochemistry,* 267, 7158-7169.

COOK, K. G., BRADFORD, A. P., YEAMAN, S. J., AITKEN, A., FEARNLEY, I. M. and WALKER, J. E. (1984) Regulation of bovine kidney branched-chain 2-oxoacid dehydrogenase complex by reversible phosphorylation. *European Journal of Biochemistry,* 145, 587-591.

COOPER, A.J., CONWAY, M. and HUTSON, S.M. (2002) A continuous 96-well plate spectrophotometric assay for branched-chain amino acidaminotransferases. *Analytical Biochemistry*, 308, 100-105.

DANSON, M. J., EISENTHAL, R., HALL, S., KESSELL, S. R. and WILLIAMS, D. L. (1984) Dihydrolipoamide dehydrogenase from halophilic archaebacteria. *Biochemical Journal,* 218, 811-818.

DANSON, M. J., FERSHT, A. R. and PERHAM, R. N. (1978) Rapid intramolar coupling of active sites in the pyruvate dehydrogenase complex of *Escherichia coli*: Mechanism for rate enhancement in a multimeric structure. *Proceedings of the National Academy of Sciences of the United States of America,* 75, 5386-5390.

DANSON, M.J., LAMBLE, H.J. and HOUGH, D.W. (2006) Central metabolism in the Archaea. *Archaea: Molecular Biology.* Editor: R. Cavicchioli. ASM Press. *In press.*

DANSON, M. J., MORGAN, D.J., JEFFRIES, A.C., HOUGH, D. W. and DYALL-SMITH, M. L. (2004) Multienzyme complexes in the Archaea: predictions from genome sequences. In *Halophilic Microorganisms.* Springer-Verlag.

DANSON, M. J., MCQUATTIE, A. and STEVENSON, K. J. (1986) Dihydrolipoamide dehydrogenase from halophilic archaebacteria - purification and properties of the enzyme from *Halobacterium halobium. Biochemistry,* 25, 3880-3884.

DARLAND, G., BROCK, T. D., SAMSONOFF, W. and CONTI, S. F. (1970) A thermophilic acidophilic mycoplasma isolated from a coal refuse pile. *Science,* 170, 1416-1418.

DAVIE, J. R., WYNN, R. M., COX, R. P. and CHUANG, D. T. (1992) Expression and assembly of a functional E1 component ($\alpha 2\beta 2$) of mammalian branched-chain alpha-ketoacid dehydrogenase complex in *Escherichia coli. Journal of Biological Chemistry,* 267, 16601-16606.

DEMELER, B., SABER, H. and HANSEN, J.C. (1997) Identification and interpretation of complexity in sedimentation velocity boundaries, *Biophysical Journal. 72* 397–407

DENOYA, C. D., FEDECHKO, R. W., HAFNER, E. W., MCARTHUR, H. A. I., MORGENSTERN, M. R., SKINNER, D. D., STUTZMANENGWALL, K., WAX, R. G. and WERNAU, W. C. (1995) A 2nd branched-chain alpha-keto acid dehydrogenase gene-cluster (Bkd) From *Streptomyces avermitilis* - its relationship to avermectin biosynthesis and the construction of a Bkd mutant suitable for the production of novel antiparasitic avermectins. *Journal of Bacteriology,* 177, 3504-3511.

DOWNARD, J. and TOAL, D. (1995) Branched-chain fatty acids: the case for a novel form of cell-cell signaling during *Myxococcus xanthus* development. *Molecular Microbiology.* 16, 171-175.

EISENTHAL, R. and CORNISH-BOWDEN, A. (1974) The Direct Linear Plot. A new graphical procedure for estimating enzyme kinetic parameters. *Biochemical Journal.* 139, 715-720.

FISHER, C., CHUANG, J., GRIFFIN, T., LAU, K., COX, R. and CHUANG, D. (1989) Molecular phenotypes in cultured maple syrup urine disease cells. Complete E1 alpha cDNA sequence and mRNA and subunit contents of the human branched chain alpha-keto acid dehydrogenase complex. *Journal of Biological Chemistry,* 264, 3448-3453.

FORTERRE, P., BROCHIER, C. and PHILIPPE, H. (2002) Evolution of the archaea. *Theoretical Population Biology,* 61, 409-422.

FRANK, R. A. W., TITMAN, C. M., PRATAP, J. V., LUISI, B. F. and PERHAM, R. N. (2004) A molecular switch and proton wire synchronize the active sites in thiamine enzymes. *Science,* 306, 872-876.

FRIES, M., CHAUHAN, H. J., DOMINGO, G. J., JUNG, H. I. and PERHAM, R. N. (2003) Site-directed mutagenesis of a loop at the active site of E1 (alpha(2)beta(2)) of the pyruvate dehydrogenase complex - A possible common sequence motif. *European Journal of Biochemistry,* 270, 861-870.

FRIES, M., JUNG, H. I. and PERHAM, R. N. (2003) Reaction mechanism of the heterotetrameric ($\alpha_2\beta_2$) E1 component of 2-oxo acid dehydrogenase multienzyme complexes. *Biochemistry,* 42, 6996-7002.

GASTEIGER, E., HOOGLAND, C., GATTIKER, A., DUVAUD, S., WILKINS, M.R., APPEL, R.D., BAIROCH, A. (2005) Protein identification and analysis tools on the ExPASy server. In *The Proteomics Handbook.* Walker J.M. ed., 571-607, Humana Press.

HAWKINS, C. F., BORGES, A. and PERHAM, R. N. (1989) A common structural motif in thiamin pyrophosphate-binding enzymes. *FEBS Letters,* 255, 77-82.

HEATH, C., JEFFRIES, A. C., HOUGH, D. W. and DANSON, M. J. (2004) Discovery of the catalytic function of a putative 2-oxoacid dehydrogenase multienzyme complex in the thermophilic archaeon *Thermoplasma acidophilum. FEBS Letters,* 577, 523-527.

HENDERSON, C. E. and PERHAM, R. N. (1980) Purification of the pyruvate-dehydrogenase multi-enzyme complex of *Bacillus stearothermophilus* and resolution of its 4 component polypeptides. *Biochemical Journal,* 189, 161-172.

HENGEVELD, A.F., SCHOUSTRA, S.E., WESTPHAL, A.H. and DE KOK, A. (1999) Pyruvate dehydrogenase from *Azotobacter vinelandii.* Properties of the N-terminally truncated enzyme. *European Journal of Biochemistry,* 265, 1098-1107.

HENSLEY, P. (1996) Defining the structure and stability of macromolecular assemblies in solution: The re-emergence of analytical ultracentrifugation as a practical tool. *Structure.* 4, 367-373.

HOLNESS, M. J. and SUGDEN, M. C. (2003) Regulation of pyruvate dehydrogenase complex activity by reversible phosphorylation. *Biochemical Society Transactions,* 31, 1143-1151.

HUBER, H., HOHN, M.J., RACHEL, R., FUCHS, T., WIMMER, V.C., STETTER, K.O. (2002) A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature.* 417, 27-8.

IZARD, T., AEVARSSON, A., ALLEN, M. D., WESTPHAL, A. H., PERHAM, R. N., DE KOK, A. and HOL, W. G. J. (1999) Principles of quasi-equivalence and Euclidean geometry govern the assembly of cubic and dodecahedral cores of pyruvate dehydrogenase complexes. *Proceedings of the National Academy of Sciences of the United States of America,* 96, 1240-1245.

JAGER, A., SAMORSKI, R., PFEIFER, F., KLUG, G. (2002) Individual gvp transcript segments in *Haloferax mediterranei* exhibit varying half-lives, which are differentially affected by salt concentration and growth phase. *Nucleic acids research,* 30, 5436-5443.

JOHNSON, Z. I. and CHISHOLM, S. W. (2004) Properties of overlapping genes are conserved across microbial genomes. *Genome Research,* 14, 2268-2272.

JOLLEY, K. A., MADDOCKS, D. G., GYLES, S. L., MULLAN, Z., TANG, S.-L., DYALL-SMITH, M. L., HOUGH, D. W. and DANSON, M. J. (2000) 2-oxoacid dehydrogenase multienzyme complexes in the halophilic *Archaea*? Gene sequences and protein structural predictions. *Microbiology,* 146, 1061-1069.

JOLLEY, K. A., RAPAPORT, E., HOUGH, D. W., DANSON, M. J., WOODS, W. G. and DYALLSMITH, M. L. (1996) Dihydrolipoamide dehydrogenase from the halophilic archaeon *Haloferax volcanii*: Homologous overexpression of the cloned gene. *Journal of Bacteriology,* 178, 3044-3048.

JONES, D. D., STOTT, K. M., HOWARD, M. J. and PERHAM, R. N. (2000) Restricted motion of the lipoyl-lysine swinging arm in the pyruvate dehydrogenase complex of *Escherichia coli. Biochemistry,* 39, 8448-8459.

JURGENS, (2002) Molecular phylogeny of Archaea in boreal forest soil, freshwater and temperate estuarine sediment. *Ph.D. thesis report.*

KAWASHIMA, T., AMANO, N., KOIKE, H., MAKINO, S., HIGUCHI, S., KAWASHIMA-OHYA, Y., WATANABE, K., YAMAZAKI, M., KANEHORI, K., KAWAMOTO, T., NUNOSHIBA, T., YAMAMOTO, Y., ARAMAKI, H., MAKINO, K. and SUZUKI, M. (2000) Archaeal adaptation to higher temperatures revealed by genomic sequence of *Thermoplasma volcanium. Proceedings of the National Academy of Sciences of the United States of America,* 97, 14257-14262.

KERSCHER, L., NOWITZKI, S. and OESTERHELT, D. (1982) Thermoacidophilic archaebacteria contain bacterial-type ferredoxins acting as electron-acceptors of 2-oxoacid-ferredoxin oxidoreductases. *European Journal of Biochemistry,* 128, 223-230.

KERSCHER, L. and OESTERHELT, D. (1981) The catalytic mechanism of 2-oxoacid-ferredoxin oxidoreductases from *Halobacterium halobium* - one-electron transfer at 2 distinct steps of the catalytic cycle. *European Journal of Biochemistry,* 116, 595-600.

KERSCHER, L. and OESTERHELT, D. (1982) Pyruvate - ferredoxin oxidoreductase New findings on an ancient enzyme. *Trends in Biochemical Sciences,* 7, 371-374.

KIM, S. and LEE, S.B. (2005) Identification and characterization of *Sulfolobus solfataricus* D-gluconate dehydratase: a key enzyme in the non-phosphorylated Entner-Doudoroff pathway. *Biochemical Journal.* 387, 271-80.

KIM, J., KIM, K.H., LEE, H.H., LEE, S.J., HA, J.Y., YOON, H.J. and SUH, S.W. (2005) Crystal structure of lipoate-protein ligase A bound with the activated intermediate: insights into interaction with the lipoyl domains. *Journal of Biological Chemistry*, 280, 38081-38089.

LAUE, T.M. and STAFFORD,W.F. (1999) Modern applications of analytical ultracentrifugation, *Annual Review of Biophysisical and Biomolecular Structure. 28*, 75–100.

LESSARD, I. A. D., DOMINGO, G. J., BORGES, A. and PERHAM, R. N. (1998) Expression of genes encoding the E2 and E3 components of the *Bacillus stearothermophilus* pyruvate dehydrogenase complex and the stoichiometry of subunit interaction in assembly in vitro. *European Journal of Biochemistry,* 258, 491-501.

LESSARD, I. A. D. and PERHAM, R. N. (1994) Expression in *Escherichia coli* of genes encoding the E1-alpha and E1-beta subunits of the pyruvate-dehydrogenase complex of *Bacillus stearothermophilus* and assembly of a functional E1 component ($\alpha 2\beta 2$) in vitro. *Journal of Biological Chemistry,* 269, 10378-10383.

LOWE, P. N., HODGSON, J. A. and PERHAM, R. N. (1983) Dual role of a single multienzyme complex in the oxidative decarboxylation of pyruvate and branched-chain 2-oxo acids in *Bacillus subtilis. Biochemical Journal,* 215, 133-140.

MADHUSUDHAN, K. T., LUO, J. H. and SOKATCH, J. R. (1999) In vitro transcriptional studies of the bkd operon of *Pseudomonas putida*: L-branched-chain amino acids and D-leucine are the inducers. *Journal of Bacteriology,* 181, 2889-2894.

MARCHLER-BAUER, A. and BRYANT, S.H. (2004) CD-Search: protein domain annotations on the fly. *Nucleic Acids Research*, Res. 32(W), 327-331.

MARSHALL, V. D. and SOKATCH, J. R. (1972) Regulation of valine catabolism in *Pseudomonas putida. Journal of Bacteriology,* 110, 1073-and.

MATTEVI, A., DE KOK, A. and PERNHAM, R. (1992) The pyruvate dehydrogenase multienzyme complex. *Current Opinion in Structural Biology,* 2, 877-887.

MATTEVI, A., SCHEIRBEEK, A.J. and HOL, W.G. (1991) Refined crystal structure of lipoamide dehydrogenase from *Azotobacter vinelandii* at 2.2 A resolution. A comparison with the structure of glutathione reductase. *Journal of Molecular Biology*, 220, 975-994.

MCMANUS, E., LUISI, B.F. and PERHAM, R.N. (2005) Structure of a putative lipoate protein ligase from *Thermoplasma acidophilum* and the mechanism of target selection for post-translational modification. *Journal of Molecular Biology* (article in press).

MILLER, J. R., BUSBY, R. W., JORDAN, S. W., CHEEK, J., HENSHAW, T. F., ASHLEY, G. W., BRODERICK, J. B., CRONAN, J. E. and MARLETTA, M. A. (2000) *Escherichia coli* LipA is a lipoyl synthase: In vitro biosynthesis of lipoylated pyruvate dehydrogenase complex from octanoyl-acyl carrier protein. *Biochemistry,* 39, 15166-15178.

MILNE, J. L. S., SHI, D., ROSENTHAL, P. B., SUNSHINE, J. S., DOMINGO, G. J., WU, X. W., BROOKS, B. R., PERHAM, R. N., HENDERSON, R. and SUBRAMANIAM, S. (2002) Molecular architecture and mechanism of an icosahedral pyruvate dehydrogenase complex: a multifunctional catalytic machine. *EMBO Journal,* 21, 5587-5598.

MORGAN, D.J. (2001) The identification and analysis of PDHC-like genes in Archaea. *Honours Degree project report.*

MORRIS, T., REED, K. and CRONAN, J., JR (1995) Lipoic acid metabolism in *Escherichia coli*: the lplA and lipB genes define redundant pathways for ligation of lipoyl groups to apoprotein. *Journal of Bacteriology,* 177, 1-10.

MORRIS, T. W., REED, K. E. and CRONAN, J. E. (1994) Identification of the gene encoding lipoate-protein ligase-a of *Escherichia coli* - Molecular-cloning and characterization of the Lpla gene and gene-product. *Journal of Biological Chemistry,* 269, 16091-16100.

NISHIZAWA, Y., YABUKI, T., FUKUDA, E. and WAKAGI, T. (2005) Gene expression and characterization of two 2-oxoacid: ferredoxin oxidoreductases from *Aeropyrum pernix* K1. *FEBS Letters,* 579, 2319-2322.

OKU, H. and KANEDA, T. (1988) Biosynthesis of branched-chain fatty-acids in *Bacillis subtilis* - a decarboxylase Is essential For branched-chain fatty-acid synthetase. *Journal of Biological Chemistry,* 263, 18386-18396.

OZAWA, Y., NAKAMURA, T., KAMATA, N., YASUJIMA, D., URUSHIYAMA, A., YAMAKURA, F., OHMORI, D. and IMAI, T. (2005) *Thermococcus profundus* 2-ketoisovalerate ferredoxin oxidoreductase, a key enzyme in the archaeal energy-producing amino acid metabolic pathway. *Journal of Biochemistry,* 137, 101-107.

PAXTON, R. and HARRIS, R. A. (1982) Isolation of rabbit liver branched-chain alpha-ketoacid dehydrogenase and regulation by phosphorylation. *Journal of Biological Chemistry,* 257, 4433-4439.

PERHAM, R. N. (1991) Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes - a paradigm in the design of a multifunctional protein. *Biochemistry,* 30, 8501-8512.

PERHAM, R.N. (2000) Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. *Annual Reviews in Biochemistry,* 69, 961-1004.

PERHAM, R. N., DAFYDD JONES, D., CHAUHAN, H. J. and HOWARD, M. J. (2002) Substrate channelling in 2-oxo acid dehydrogenase multienzyme complexes. *Biochemical Society,* 30, 47-51.

PETTIT, F. H., YEAMAN, S. J. and REED, L. J. (1978) Purification and characterization of branched-chain alpha-keto acid dehydrogenase complex of bovine kidney. *Proceedings of the National Academy of Sciences of the United States of America,* 75, 4881-4885.

PLAGA, W., LOTTSPEICH, F. and OESTERHELT, D. (1992) Improved purification, crystallization and primary structure of pyruvate ferredoxin oxidoreductase from *Halobacterium halobium European Journal of Biochemistry.* 205, 391-397.

PRATT, K. J., CARLES, C., CARNE, T. J., DANSON, M. J. and STEVENSON, K. J. (1989) Detection of bacterial lipoic acid - a modified gas-chromatographic-mass-spectrometric procedure. *Biochemical Journal,* 258, 749-754.

RAY, W.K., KEITH, S.M., DESANTIS, A.M., HUNT, J.P., LARSON, T.J., HELM, R.F., KENELLY, P.J. (2005) A phosphohexomutase from the archaeon Sulfolobus solfataricus is covalently modified by phosphorylation on serine. *Journal of Bacteriology*, 187, 4270-5.

RECHE, P. and PERHAM, R. N. (1999) Structure and selectivity in post-translational modification: attaching the biotinyl-lysine and lipoyl-lysine swinging arms in multifunctional enzymes. *EMBO Journal,* 18, 2673-2682.

REED, L. J. (1974) Multienzyme complexes. *Accounts of Chemical Research,* 7, 40-46.

REED, L. J., PETTIT, F. H., ELEY, M. H., HAMILTON, L., COLLINS, J. H. and OLIVER, R. M. (1975) Reconstitution of the *Escherichia coli* pyruvate dehydrogenase complex. *Proceedings of the National Academy of Sciences of the United States of America,* 72, 3068-3072.

RUEPP, A., GRAML, W., SANTOS-MARTINEZ, M. L., KORETLE, K. K., VOLKER, C., MEWES, H. W., FRISHMAN, D., STOCKER, S., LUPAS, A. N. and BAUMEISTER, W. (2000) The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum. Nature,* 407, 508-513.

SCHNARRENBERGER, C.and MARTIN, W. (2002) Evolution of the enzymes of the citric acid cycle and the glyoxylate cycle of higher plants. A case study of endosymbiotic gene transfer. *European Journal of Biochemistry.* 269, 868-875.

SCHREINER, M. E., FIUR, D., HOLATKOJ, J., PATEK, M. and EIKMANNS, B. J. (2005) E1 enzyme of the pyruvate dehydrogenase complex in *Corynebactetium glutamicum*: Molecular analysis of the gene and phylogenetic aspects. *Journal of Bacteriology,* 187, 6005-6018.

SCHUT, G. J., MENON, A. L. and ADAMS, M. W. W. (2001) 2-keto acid oxidoreductases from *Pyrococcus furiosus* and *Thermococcus litoralis.* In *Methods in Enzymology,* Academic Press, San Diego.

SEGERER, A., LANGWORTHY, T. A. and STETTER, K. O. (1988) *Thermoplasma acidophilum* and *Thermoplasma volcanium* Sp-Nov from Solfatara Fields. *Systematic and Applied Microbiology,* 10, 161-171.

SKINNER, D. D., MORGENSTERN, M. R., FEDECHKO, R. W. and DENOYA, C. D. (1995) Cloning and sequencing of a cluster of genes encoding branched-chain alpha-keto acid dehydrogenase from *Streptomyces Avermitilis* and the production of a functional E1[Alpha-Beta] component in *Escherichia coli. Journal of Bacteriology,* 177, 183-190.

SLUPSKA, M. M., KING, A. G., FITZ-GIBBON, S., BESEMER, J., BORODOVSKY, M. and MILLER, J. H. (2001) Leaderless transcripts of the crenarchaeal hyperthermophile *Pyrobaculum aerophilum. Journal of Molecular Biology,* 309, 347-360.

SMITH, L. D., BUNGARD, S. J., DANSON, M. J. and HOUGH, D. W. (1987) Dihydrolipoamide dehydrogenase from the thermoacidophilic Archaebacterium *Thermoplasma acidophilum. Biochemical Society Transactions,* 15, 1097-1097.

SNIJDERS, A.P., WALTHER, J., PETER, S., KINNMAN, I., DE VOS, M.G., VAN DE WERKEN, H.J., BROUNS, S.J., VAN DER OOST, J. and WRIGHT, P.C. (2006) Reconstruction of central carbon metabolism in *Sulfolobus solfataricus* using a two-dimensional gel electrophoresis map, stable isotope labeling and DNA microarray analysis. *Proteomics,* in press.

SOKATCH, J. R., MCCULLY, V. and ROBERTS, C. M. (1981) Purification of a branched-chain keto acid dehydrogenase from *Pseudomonas putida. Journal of Bacteriology,* 148, 647-652.

SYKES, P. J., BURNS, G., MENARD, J., HATTER, K. and SOKATCH, J. R. (1987) Molecular-cloning of genes encoding branched-chain keto acid dehydrogenase of *Pseudomonas putida. Journal of Bacteriology,* 169, 1619-1625.

THOMPSON, J.D., HIGGINS,D.G and GIBSON, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22, 4673-4680.

VETTAKKORUMAKANKAV, N. N., DANSON, M.J., HOUGH, D.W., STEVENSON, K. J., DAVISON, M., YOUNG, J. (1992) Dihydrolipoamide dehydrogenase from the halophilic archaebacterium *Haloferax volcanii*: characterization and N-terminal sequence. *Biochemistry and Cell Biology*, 70, 70-75.

VETTAKKORUMAKANKAV, N. N. and STEVENSON, K. J. (1992) Dihydrolipoamide dehydrogenase from *Haloferax-volcanii* - gene cloning, complete primary structure, and comparison to other dihydrolipoamide dehydrogenases. *Biochemistry and Cell Biology,* 70, 656-663.

WALLIS, N. G. and PERHAM, R. N. (1994) Structural dependence of post-translational modification and reductive acetylation of the lipoyl domain of the pyruvate dehydrogenase multienzyme complex. *Journal of Molecular Biology,* 236, 209-216.

WANNER, C. and SOPPA, J. (2002) Functional role for a 2-oxoacid dehydrogenase in the halophilic archaeon *Haloferax volcanii. Journal of Bacteriology,* 184, 3114-3121.

WARD, D. E., ROSS, R. P., VAN DER WEIJDEN, C. C., SNOEP, J. L. and CLAIBORNE, A. (1999) Catabolism of branched-chain alpha-keto acids in *Enterococcus faecalis*: the bkd gene cluster, enzymes, and metabolic route. *Journal of Bacteriology,* 181, 5433-5442.

WARD, D. E., VAN DER WEIJDEN, C. C., VAN DER MERWE, M. J., WESTERHOFF, H. V., CLAIBORNE, A. and SNOEP, J. L. (2000) Branched-chain alpha-keto acid catabolism via the gene products of the bkd operon in *Enterococcus faecalis*: a new, secreted metabolite serving as a temporary redox sink. *Journal of Bacteriology,* 182, 3239-3246.

WITZMANN, S. and BISSWANGER, H. (1998) The pyruvate dehydrogenase complex from thermophilic organisms: thermal stability and re-association from the enzyme components. *Biochimica and Biophysica Acta,* 1385.

WOESE, C. R., KANDLER, O. and WHEELIS, M. L. (1990) Towards a natural system of organisms - proposal for the Domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America,* 87, 4576-4579.

WYNN, R. M., CHUANG, J. L., DAVIE, J. R., FISHER, C. W., HALE, M. A., COX, R. P. and CHUANG, D. T. (1992) Cloning and expression in *Escherichia coli* of mature E1-beta- subunit of bovine mitochondrial branched-chain alpha-keto acid dehydrogenase complex - mapping of the E1-beta-binding region on E2. *Journal of Biological Chemistry,* 267, 1881-1887.

WYNN, R. M., KATO, M., MACHIUS, M., CHUANG, J. L., JUN, L., TOMCHICK, D. R. and CHUANG, D. T. (2004) Molecular mechanism for regulation of the human mitochondrial branched-chain alpha-ketoacid dehydrogenase complex by phosphorylation. *Structure,* 12, 2185-2196.

YASUDA, M., OYAIZU, H., YAMAGISHI, A. and OSHIMA, T. (1995) Morphological variation of new *Thermoplasma acidophilum* isolates from Japanese hot-springs. *Applied and Environmental Microbiology,* 61, 3482-3485.

ZAIGLER, A., SCHUSTER, S. C. and SOPPA, J. (2003) Construction and usage of a onefold-coverage shotgun DNA microarray to characterize the metabolism of the archaeon *Haloferax volcanii. Molecular Microbiology,* 48, 1089-1105.

ZHANG, B., HEALY, P., ZHAO, Y., CRABB, D. and HARRIS, R. (1990) Premature translation termination of the pre-E1 alpha subunit of the branched chain alpha-ketoacid dehydrogenase as a cause of maple syrup urine disease in Polled Hereford calves. *Journal of Biological Chemistry,* 265, 2425-2427.

APPENDIX: Publication

FEBS 28987 FEBS Letters 577 (2004) 523–527

Discovery of the catalytic function of a putative 2-oxoacid dehydrogenase multienzyme complex in the thermophilic archaeon *Thermoplasma acidophilum*

Caroline Heath, Alex C. Jeffries, David W. Hough, Michael J. Danson*

Centre for Extremophile Research, Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK

Received 3 September 2004; revised 20 October 2004; accepted 21 October 2004

Available online 3 November 2004

Edited by Stuart Ferguson

Abstract Those aerobic archaea whose genomes have been sequenced possess a single 4-gene operon that, by sequence comparisons with Bacteria and Eukarya, appears to encode the three component enzymes of a 2-oxoacid dehydrogenase multienzyme complex. However, no catalytic activity of any such complex has ever been detected in the Archaea. In the current paper, we have cloned and expressed the first two genes of this operon from the thermophilic archaeon, *Thermoplasma acidophilum*. We demonstrate that the protein products form an $\alpha_2\beta_2$ hetero-tetramer possessing the decarboxylase catalytic activity characteristic of the first component enzyme of a branched-chain 2-oxoacid dehydrogenase multienzyme complex. This represents the first report of the catalytic function of these putative archaeal multienzyme complexes.

Keywords: Archaea; 2-Oxoacid dehydrogenase; Multienzyme complex; Thermophile; Metabolism

1. Introduction

In aerobic bacteria and eukaryotes, a family of 2-oxoacid dehydrogenase multienzyme complexes (OADHCs) functions in the pathways of central metabolism. The complexes are responsible for the oxidative decarboxylation of 2-oxoacids to their corresponding acyl-CoAs (Fig. 1). Members of the family include the pyruvate dehydrogenase complex (PDHC), which catalyses the conversion of pyruvate to acetyl-CoA and so links glycolysis and the citric acid cycle; the 2-oxoglutarate dehydrogenase complex (OGDHC), which catalyses the conversion of 2-oxoglutarate to succinyl-CoA within the citric acid cycle; and the branched-chain 2-oxoacid dehydrogenase complex (BCOADHC), which oxidatively decarboxylates the branched-chain 2-oxoacids produced by the transamination of valine, leucine and isoleucine.

The complexes comprise multiple copies of three component enzymes: 2-oxoacid decarboxylase (E1), dihydrolipoyl acyltransferase (E2) and dihydrolipoamide dehydrogenase (E3) [1–3]. E2 forms the structural core of the complex, with multiple polypeptide chains associating into octahedral (24-mer) or icosahedral (60-mer) configurations, depending on the particular complex and the source organism [2,4]. E1 and E3 bind non-covalently to the E2 core. E1 may occur as a homodimer or as an $\alpha_2\beta_2$ hetero-tetramer, depending upon the source and the type of complex, although in all cases E3 is a dimer of identical subunits.

E2 also forms the catalytic core of the complex (Fig. 1): a lipoyl moiety, covalently attached to a lysine residue in the lipoyl domain, serves as a swinging arm, connecting the active sites of each enzyme and channelling substrate through the complex [3]. Thus, E1 catalyses the thiamine pyrophosphate (TPP)-dependent oxidative decarboxylation of the 2-oxoacid and the transfer of the resulting acyl group to the lipoic acid of E2. E2 then transfers the acyl-group to coenzyme-A, after which E3 serves to reoxidise the dihydrolipoyl moiety. It does so by the reduction of the non-covalently bound co-factor FAD, in conjunction with a protein disulfide bond and an amino acid base, which are themselves then reoxidised by NAD^+, forming NADH.

No OADHC activity has ever been detected in the Archaea [5]; instead, the oxidation of 2-oxoacids is catalysed by an unrelated and structurally more simple family of 2-oxoacid ferredoxin oxidoreductases (FORs). This comprises the pyruvate FOR, the 2-oxoglutarate FOR and the 2-oxoisovalerate FOR, which catabolise pyruvate, 2-oxoglutarate and the branched-chain 2-oxoacids, respectively [6–8]. The pyruvate FOR from *Halobacterium halobium* has an $\alpha_2\beta_2$ structure [9], but in the thermophilic archaea these enzymes generally contain four different subunits and occur as octamers ($\alpha_2\beta_2\gamma_2\delta_2$) [8]. Importantly in the current context, the FOR's catalytic reaction does not involve a lipoic acid moiety or NAD^+; rather, the acyl-moiety formed on decarboxylation of the 2-oxoacid is handed on direct to coenzyme-A, and the reducing equivalents to ferredoxin via the enzyme's iron-sulfur centre [6].

However, there is growing evidence to suggest that some archaea may possess an OADHC (reviewed in [10]). Whilst no whole complex activity has been found, E3 activity and lipoic acid have been detected in halophilic archaea [11,12] and an operon containing four genes, whose predicted protein

* Corresponding author. Fax: +44-1225-386779.
E-mail address: M.J.Danson@bath.ac.uk (M.J. Danson).

Abbreviations: BCOADHC, branched-chain 2-oxoacid dehydrogenase complex; DCPIP, 2,6-dichlorophenolindophenol; E1, 2-oxoacid decarboxylase; E2, dihydrolipoyl acyl-transferase; E3, dihydrolipoamide dehydrogenase; FOR, ferredoxin oxidoreductases; M_r, relative molecular mass; OADHC, 2-oxoacid dehydrogenase complexes; OGDHC, 2-oxoglutarate dehydrogenase complex; PDHC, pyruvate dehydrogenase complex; TPP, thiamine pyrophosphate

0014-5793/$22.00 © 2004 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/j.febslet.2004.10.058

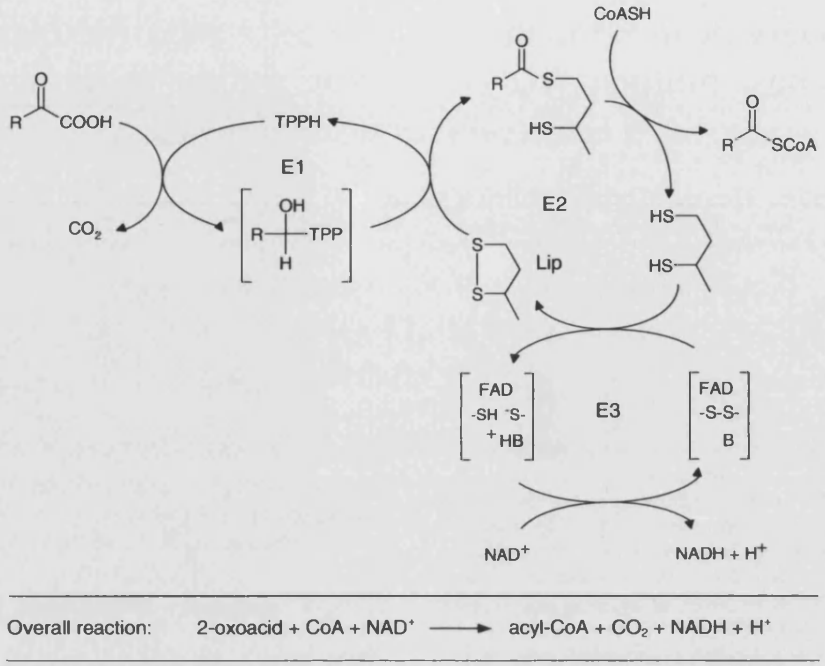


Fig. 1. The general reaction catabolised by OADHCs. The complex mechanism is described in Section 1. E1 (2-oxoacid decarboxylase); E2 (dihydrolipoyl acyl-transferase); E3 (dihydrolipoamide dehydrogenase). Symbols: B (a histidine base on E3); Lip (enzyme-bound lipoic acid); CoASH (coenzyme-A); TPP-H (thiamine pyrophosphate); FAD (flavin adenine dinucleotide); NAD (nicotinamide adenine dinucleotide).

sequences show significant identity to OADHC components E1α, E1β, E2 and E3 of bacteria and eukaryotes, was subsequently found in *Haloferax volcanii* [13]. E3 activity has also been detected in cell extracts of *Thermoplasma acidophilum* [14]. The recent sequencing of a number of archaeal genomes has revealed the presence of putative OADHC-encoding operons in the aerobic archaea: *H. halobium*, *T. acidophilum*, *Aeropyrum pernix*, *Pyrobaculum aerophilum*, and *Sulfolobus solfataricus*. To date, there is no experimental evidence to suggest that a functional complex is assembled from these operons, and E3-knockout studies in *Hfx. volcanii* failed to elucidate any metabolic function [15].

In the current paper, to elucidate the function of this putative OADHC in the Archaea, we have cloned and expressed in *Escherichia coli* the E1α and E1β genes from the archaeon, *T. acidophilum*, which is a thermoacidophile that grows optimally at 59 °C and pH 2.0 [16]. Enzyme assays of the purified recombinant E1 enzyme show it to be an active branched-chain 2-oxoacid decarboxylase component, suggesting that the operon in this archaeon encodes a functional BCOADHC.

2. Materials and methods

2.1. Bioinformatics

The putative OADHC operon was identified in the *T. acidophilum* DSM1728 genome from the ENTREZ Nucleotides database (http://www3.ncbi.nlm.nih.gov). The features of this operon and the arrangement of the four genes [E1α: Ta1438; E1β: Ta1437; E2: Ta1436; and E3: Ta1435] are illustrated diagrammatically in Fig. 2.

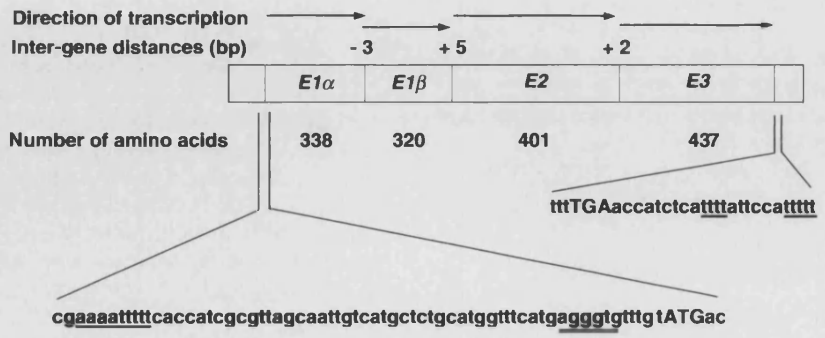


Fig. 2. The arrangement, inter-gene distances (base pairs, bp) and proposed direction of transcription (→) of the four open reading frames constituting the proposed *T. acidophilum* OADHC operon. The corresponding number of amino acids of the protein products is given. The DNA sequence at the 5′ end of the E1α gene, plus the upstream region, is shown to illustrate a potential promoter sequence (underlined), a potential Shine–Dalgarno sequence (doubly underlined) and the ATG start codon. The 3′ TGA stop codon and the proposed transcriptional stop signals (underlined) are also shown.

2.2. Gene cloning

T. acidophilum DSM 1728 was grown at 59 °C in DSM medium 158 and, after cell lysis in 50 mM Tris buffer (pH 7.5), genomic DNA was extracted according to [17]. The genes encoding E1α and E1β were individually PCR-amplified from this *T. acidophilum* genomic DNA, using primers that introduced restriction sites to the 5′ and 3′ ends of the gene products (*Xho*I and *Bam*HI for E1α, *Nhe*I and *Bam*HI for E1β). These primers also permitted the introduction of an N-terminal His tag into the protein products when expressed in the pET vector system (Novagen). PCR-amplification was carried out using *Vent* polymerase (New England Biolabs), followed by A-tailing with *Taq* polymerase (Promega). Amplified DNA fragments were purified by electrophoresis in a 0.8% (w/v) agarose gel and extracted using the Qiaex II Gel Extraction kit (Qiagen). Fragments were separately ligated into the pGEM-T (Promega) intermediate cloning vector, from which the E1α gene fragment was excised and cloned into pET-19b (Amp^r), and the E1β fragment into pET-28a (Kan^r), using T4 DNA ligase, generating the recombinant expression vectors pET-19b-E1α and pET-28a-E1β.

2.3. Expression of gene products

E. coli BL21(DE3)pLysS cells were separately heat-shock transformed with the pET-19b-E1α and pET-28a-E1β constructs. For co-transformation with both plasmids, cells were first transformed with plasmid pET-19b-E1α and selected by growth on LB agar containing carbenicillin (50 μg/ml); transformants were then made competent by treatment with calcium chloride and subjected to a second transformation, resulting in pET-19b-E1α-pET-28a-E1β co-transformants, selected for by growth on media containing both carbenicillin (50 μg/ml) and kanamycin (30 μg/ml).

Single-plasmid transformants were grown in LB media containing chloramphenicol (34 μg/ml) and either carbenicillin (50 μg/ml) or kanamycin (30 μg/ml). Following induction with 1 mM isopropyl-β-D-thiogalactose at $OD_{600\ nm} \approx 0.6$, cells were grown for a further 4 h at 37 °C with shaking. Co-transformants were grown in media containing all three antibiotics and the proteins expressed similarly.

2.4. Purification of recombinant protein

Co-expressed recombinant E1 proteins were purified by His-Bind Resin (Novagen) chromatography and dialysed into 20 mM Tris buffer (pH 9.3), containing 10% (v/v) glycerol. After 24 h, precipitated protein was isolated by centrifugation (16000 × *g* for 5 min), resuspended in the same dialysis buffer and heated for 20 min at 55 °C. After this treatment, remaining insoluble precipitate was removed by a similar centrifugation and discarded. Purity of E1 in the retained soluble fraction was examined by SDS–PAGE on a 10% (w/v) polyacrylamide gel.

2.5. Gel filtration

The size of the recombinant E1αβ protein was estimated by analytical gel filtration on an Amersham Biosciences Äkta FPLC system, using a Superdex 200 10/300GL column equilibrated with 20 mM sodium phosphate (pH 7.0), 2 mM $MgCl_2$ and 100 mM NaCl. Protein standards were: β-amylase ($M_r = 200000$), alcohol dehydrogenase (150000), bovine serum albumin (66000), carbonic anhydrase (29000) and cytochrome *C* (12400).

2.6. Densitometry

The E1α and E1β protein bands on SDS–PAGE gel were analysed densitometrically. The Coomassie blue stained gel was scanned using a Fujifilm FLA-5000 phosphorimager at 473 nm, with an LPG filter (for general-purpose excitation fluorescent digitising), and the scanned image was then analysed using the AIDA 2D Densitometry programme (Raytest, Straubenhardt, Germany).

2.7. Enzyme assay

E1 enzymic activity was assayed spectrophotometrically by following the 2-oxoacid dependent reduction of 2,6-dichlorophenolindophenol (DCPIP) at 595 nm [18]. Assays were carried out at 55 °C in 20 mM potassium phosphate (pH 7.0), 2 mM $MgCl_2$ and 0.2 mM TPP. Buffer and recombinant E1α enzyme were pre-incubated at 55 °C for 10 min; 50 μM DCPIP was then added and the assay started by the addition of the 2-oxoacid substrate [either pyruvate, 2-oxoglutarate, 4-methyl-2-oxopentanoate, 3-methyl-2-oxopentanoate or 3-methyl-2-oxobutanoate (Sigma–Aldrich)]. Kinetic parameters were determined by the direct linear method of Eisenthal and Cornish-Bowden [19].

3. Results

3.1. E1 expression and purification

All expression experiments used the *E. coli* BL21(DE3)-pLysS cell line and were performed at 37 °C. Two different methods were used in the attempt to obtain assembled E1 protein. The first was the individual expression of E1α and E1β, to be followed by in vitro mixing of the purified subunits, whereas the second method involved co-expression, which offered the potential of in vivo assembly of the subunits in the same cellular compartment. In the latter, cells were co-transformed with the recombinant pET-19b-E1α and pET-28a-E1β plasmids, resulting in *E. coli* pET-19b-E1α/pET-28a-E1β co-transformants.

High levels of soluble E1α were obtained by individual expression of this subunit, although E1β expression was not detectable under the same conditions (results not shown). As co-expression experiments were successful in yielding soluble E1α and E1β subunits, further manipulations to the individual expression of E1β were not attempted.

Following dialysis of the His-purified E1αβ co-expressed protein into 20 mM Tris buffer (pH 9.3) containing 10% (v/v) glycerol to remove the imidazole, the majority of protein precipitated. Resuspension of precipitated protein at 55 °C in the dialysis buffer resulted in a highly purified E1 sample with an approximately equal ratio of α to β polypeptides. The M_r values of the α and β polypeptides were determined by SDS–PAGE to be 43000 and 37000, respectively (Fig. 3), which compares favourably with predicted values of 41000 and 37000 from the published gene sequences. Analytical gel filtration revealed a single peak of around $M_r = 165000$ that was enzymically active in the E1 assay. SDS–PAGE of the peak fraction after gel filtration and subsequent densitometric analysis showed a 1:1 stoichiometry of the α and β polypeptides. Taken together, these data suggest an $\alpha_2\beta_2$ structure for the active E1 enzyme.

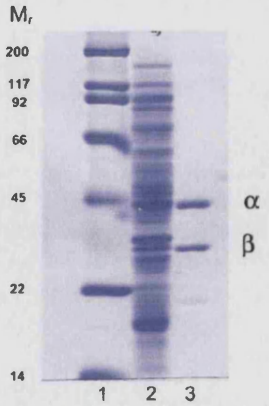


Fig. 3. SDS–PAGE of purified *T. acidophilum* recombinant E1. Lane 1, standard protein markers (M_r values given in kDa); lane 2, soluble cell extract; lane 3, enzyme after His-Bind Resin chromatography (E1α and E1β polypeptides are indicated).

Table 1
Kinetic parameters determined for purified recombinant E1 from *T. acidophilum* at 55 °C with branched-chain 2-oxoacids, pyruvate and 2-oxoglutarate as potential substrates

2-Oxoacid substrate	V_{max} (U mg^{-1} protein)	K_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (mM^{-1} s^{-1})
3-Methyl-2-oxopentanoate	0.38	0.50	0.1	5
4-Methyl-2-oxopentanoate	0.13	0.17	0.042	4
3-Methyl-2-oxobutanoate	0.23	0.30	0.027	11
Pyruvate	0.07	0.09	0.6	0.15
2-Oxoglutarate	0	0	0	0

1 Unit of enzyme activity (U) is defined as 1 μmol DCPIP reduced per min.
Values of k_{cat} are calculated per αβ dimer of the $\alpha_2\beta_2$ active enzyme.

3.2. E1 activity

The recombinant E1 enzyme was incubated with TPP for 10 min prior to assay, a lag in the production of product being observed if the assay was initiated without prior incubation. However, increasing the pre-incubation period to 4 h resulted in no increased E1 activity over that obtained with the 10 min incubation. E1 enzymic activity was detected with the 2-oxoacids 3-methyl-2-oxopentanoate, 4-methyl-2-oxopentanoate, 3-methyl-2-oxobutanoate and pyruvate, but not with 2-oxoglutarate. The calculated kinetic parameters with these substrates are given in Table 1.

4. Discussion

The discovery in the aerobic archaea of an operon that would appear to encode the components of a OADHC was unexpected because the catalytic activity of such multienzyme systems has never been detected in any archaeon [10]. Also, the presence throughout the Archaea of catalytically active 2-oxoacid FORs that catalyse the same chemical conversions as the OADHCs would argue against a need for the latter multienzyme systems in these organisms.

However, the detection of enzyme activity for the E3 component, dihydrolipoamide dehydrogenase, in cell extracts of the extreme halophiles [11] and in *T. acidophilum* [14], suggests that at least this gene within the operon is functional. Furthermore, sequence analyses and secondary-structural predictions of the putative E1 and E2 enzymes (see [10,13] for programmes used) support the possibility that these are also functional proteins. Hawkins et al. [20] describe a common sequence motif found on the E1α of TPP-binding enzymes, and propose that this element is associated with TPP binding. The sequence is generally around 30 amino acids in length, with highly conserved GDG and NN residues at the N- and C-termini, respectively. Internally, there is an E or D residue that is usually conserved, and A and P residues that are generally conserved. A cluster of hydrophobic residues immediately precedes the NN [20]. As shown in Fig. 4, the motif has been located in the E1α of *T. acidophilum*. In all the bacterial and eukaryotic OADHCs, the E2 polypeptide chain has three different domains: one to three copies of an N-terminal lipoyl domain, an E1–E3 subunit-binding domain, and a catalytic domain. Structural predictions of the proposed protein product of the E2 gene in *Hfx. volcanii* and *T. acidophilum* show that this characteristic domain structure and catalytically important residues, including the lysine to which the lipoic acid is covalently bound, are conserved [10].

Given the above experimental results, and since genome analyses argue strongly against the presence of non-functional ORFs in prokaryotes [21], the work reported in this paper set out to discover the possible catalytic function of the putative OADHCs in Archaea. As the substrate specificity of OADHCs is determined by the E1 component, we chose to express the two genes that, by sequence comparisons, were predicted to code for the α and β polypeptides of the E1 decarboxylase component of the complex from the thermophilic archaeon, *T. acidophilum*. Co-expression of the putative E1α and E1β genes in *E. coli* was necessary to obtain the β protein in soluble form, suggesting that the α and β components associated with each other. This association was supported by gel filtration and SDS–PAGE, an $\alpha_2\beta_2$ enzyme being produced.

The recombinant E1 enzyme was found to possess catalytic activity with the branched-chain 2-oxoacids derived from the transamination of the amino acids, valine, leucine and isoleucine. It also exhibited activity with pyruvate as substrate, but the catalytic efficiency, defined as k_{cat}/K_m, was 25–70 times lower than with the branched-chain 2-oxoacids. No activity could be found with 2-oxoglutarate. Thus, the enzyme may be the first component of a BCOADHC, involved in the catabolism of valine, leucine and isoleucine, the acyl-CoA products then being fed into the citric acid cycle. The E1$\alpha_2\beta_2$ structure and the catalytic properties are similar to those observed in the bacterium *Bacillus subtilis*, where it is proposed that a single multienzyme complex performs an economic 'dual role' of catalysing the oxidative decarboxylation of both branched-chain 2-oxoacids and pyruvate [22]. However, in that case, the activity detected with pyruvate was approximately 50 times greater than with the branched-chain 2-oxoacids.

Finally, it should be noted that the catalytic activity of the recombinant *T. acidophilum* E1 [$k_{cat} = 0.17$–0.50 s^{-1}; Table 1] is similar to that reported for the recombinant E1$\alpha_2\beta_2$

```
T. acidophilum  (147)       fgdggtstpdfhaamnfaavfdlpvvflcennq
                             ******  **    ***  *  *  *   **
B. stearothermophilus (172) tgdggtsqgdfyeginfagafkapaifvvqnnr
```

Fig. 4. Alignment of a partial amino acid sequence of *T. acidophilum* E1α with the putative TPP-binding motif of the E1α protein of the PDHC of *Bacillus stearothermophilus*. Highly conserved regions associated with a putative TPP-binding motif in TPP-binding enzymes, as described in [20], are indicated in bold type. Hydrophobic amino acid clusters are underlined. Identities are indicated by an asterisk. Bracketed numbers give the first amino acid position in the E1α polypeptide.

[$k_{cat} = 0.47\ s^{-1}$] from the *Bacillus stearothermophilus* PDHC, using the same DCPIP assay but with pyruvate as substrate [18].

In conclusion, our data suggest that *T. acidophilum* possesses a functional E1 that accepts branched-chain 2-oxoacids as substrate. This is the first such activity to be reported in the Archaea and, with E3 activity already detected in cell extracts of this organism, is highly suggestive of a functional BCOADHC in this evolutionary domain of organisms. Thus, it is reasonable to propose that a OADHC was present in the common ancestor to the Bacteria and the Archaea, and that it has been retained in aerobic members of each domain. The 2-oxoacid FORs are also ancient enzymes but, whereas they are found in the anaerobic bacteria, they are functional in both aerobic and anaerobic archaea [7,8]. Furthermore, in the bacterial domain, the ancestral OADHC appears to have diversified to give several complexes of different substrate specificities (PDHC, OGDHC and the BCOADHC). We can find only a single complete OADHC operon in each aerobic archaeon, although a second but partial operon has been discovered in *Hfx. volcanii* [23]. This second operon encodes E1α and E1β proteins, and an unattached lipoyl domain; whilst no enzyme activity could be assigned to these products, mutation-complementation experiments suggest that they are functional during nitrate-respirative growth on Casamino acids.

It remains to be established whether the complexes from the various archaea all have the same substrate specificity, if that particular activity is always duplicated by the corresponding FOR in each organism, and how the respective activities are coordinately regulated. Structurally, the biophysical nature of the archaeal multienzyme complex is a key question, and recombinant E2 and E3 components of *T. acidophilum* must be now generated, and the ambitious goal of complex assembly achieved. The identification of the catalytic activity of the E1 component in this current paper is an important step to enabling these questions to be investigated.

Acknowledgements: C.H. gratefully acknowledges the receipt of a Research Studentship from the Biotechnology and Biological Sciences Research Council, UK. We also thank Simon Willies, University of Bath, for skilled experimental assistance.

References

[1] Perham, R.N. (1991) Biochemistry 30, 8501–8512.
[2] Perham, R.N. (2000) Ann. Rev. Biochem. 69, 961–1004.
[3] Perham, R.N., Jones, D.D., Chauhan, H.J. and Howard, M.J. (2002) Biochem. Soc. Trans. 30, 47–51.
[4] Izard, T., Aevarsson, A., Allen, M.D., Westphal, A.H., Perham, R.N., de Kok, A. and Hol, W.G.J. (1999) Proc. Natl. Acad. Sci. USA 96, 1240–1245.
[5] Danson, M.J. (1993) New Comp. Biochemistry [The Biochemistry of Archaea] 26, 1–24.
[6] Kerscher, L. and Oesterhelt, D. (1981) Eur. J. Biochem. 116, 595–600.
[7] Kerscher, L. and Oesterhelt, D. (1982) Trends Biochem. Sci. 7, 371–374.
[8] Schut, G.J., Menon, A.L. and Adams, M.W.W. (2001) Methods Enzymol. 331, 144–158.
[9] Plaga, W., Lottspeich, F. and Oesterhelt, D. (1992) Eur. J. Biochem. 205, 391–397.
[10] Danson, M.J., Morgan, D.J., Jeffries, A.C., Hough, D.W. and Dyall-Smith, M.L. (2004) in: Halophilic Microorganisms (Ventosa, A., Ed.), pp. 177–191, Springer, Berlin.
[11] Danson, M.J., Eisenthal, R., Hall, S., Kessell, S.R. and Williams, D.L. (1984) Biochem. J. 218, 811–818.
[12] Pratt, K.J., Carles, C., Carne, T.J., Danson, M.J. and Stevenson, K.J. (1989) Biochem. J. 258, 749–754.
[13] Jolley, K.A., Maddocks, D.G., Gyles, S.L., Mullan, Z., Tang, S.-L., Dyall-Smith, M.L., Hough, D.W. and Danson, M.J. (2000) Microbiology 146, 1061–1069.
[14] Smith, L.D., Bungard, S.J., Danson, M.J. and Hough, D.W. (1987) Biochem. Soc. Trans. 15, 1097.
[15] Jolley, K.A., Rapaport, E., Hough, D.W., Danson, M.J., Woods, W.G. and Dyall-Smith, M.L. (1996) J. Bacteriol. 178, 3044–3048.
[16] Darlang, G., Brock, T.D., Samsonoff, W. and Conti, S.F. (1970) Science 170, 1416–1418.
[17] Sambrook, J. and Russell, D.W. (2001) Molecular Cloning: A Laboratory Manual, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.
[18] Lessard, I.A.D. and Perham, R.N. (1994) J. Biol. Chem. 269, 10378–10383.
[19] Eisenthal, R. and Cornish-Bowden, A. (1974) Biochem. J. 139, 715–720.
[20] Hawkins, C.F., Borges, A. and Perham, R.N. (1989) FEBS Lett. 225, 77–82.
[21] Almi, J-P., Poirot, O., Lopez, F. and Claverie, J.-M. (2000) Genome Res. 10, 959–966.
[22] Lowe, P.N., Hodgson, J.A. and Perham, R.N. (1983) Biochem. J. 215, 133–140.
[23] Wanner, C. and Soppa, J. (2002) J. Bacteriol. 184, 3114–3121.